GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 00:36:13 ; Search time 456 Seconds
(without alignments)
9951.264 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2015	100.0	2015	24	ABK83939	Human cDNA differe
2	2015	100.0	2015	24	ABL66673	Lung cancer relate
3	1552	77.0	1926	24	ABK83940	Human cDNA differe
4	183	9.1	183	24	ABL61214	Human nucleotide f
5	181	9.0	1416	24	ABL61215	Rat/human fusion c
6	181	9.0	1542	24	ABL61216	Rat/human fusion c
7	169	8.4	369	16	AAT19957	Human gene signatu
8	133	6.6	171	22	ABA50558	Human breast cell
9	133	6.6	171	22	ABA68516	Human foetal liver
10	133	6.6	171	22	ABA35497	Probe #13963 for g
11	133	6.6	171	22	AAK16884	Human brain expres
12	133	6.6	171	22	AAK42654	Human bone marrow
13	133	6.6	171	22	AAI23408	Probe #13341 for g
14	133	6.6	171	22	AAI48728	Probe #17414 used
15	133	6.6	171	22	AAI09035	Probe #9026 used t
16	133	6.6	171	24	ABS16706	Human genome-deriv
17	113	5.6	415	22	ABA45430	Human breast cell
18	113	5.6	415	22	ABA55928	Human foetal liver
19	113	5.6	415	22	ABA25595	Probe #4061 for ge
20	113	5.6	415	22	AAK04142	Human brain expres
21	113	5.6	415	22	AAK29623	Human bone marrow
22	113	5.6	415	22	AAI14202	Probe #4135 for ge
23	113	5.6	415	22	AAI35583	Probe #4269 used t
24	113	5.6	415	22	AAI04039	Probe #4030 used t
25	113	5.6	415	24	ABS04179	Human genome-deriv
26	112	5.6	1592	20	AAZ27241	Human secreted pro
27	78	3.9	409	22	AAH99174	Human protein enco
28	77	3.8	334	21	AAA52650	Eosinophil activat
29	68	3.4	1911	24	ABK63704	Rat sequence diffe
30	66	3.3	274	22	AAK68573	Human immune/haema
c 31	65	3.2	1926	24	ABK83940	Human cDNA differe
32	31	1.5	31	22	AAI30734	Human single nucle
33	31	1.5	31	22	AAI30735	Human single nucle
34	31	1.5	31	22	AAI30736	Human single nucle
35	31	1.5	31	22	AAI30737	Human single nucle
36	31	1.5	31	22	AAI30738	Human single nucle
37	28	1.4	2298	24	ABK83935	Human cDNA differe
c 38	27	1.3	33	22	AAH41498	Human tyrosine kin
39	26	1.3	32	22	AAH41491	Human tyrosine kin
c 40	26	1.3	32	22	AAH41492	Human tyrosine kin
41	26	1.3	1602	14	AAQ46687	Chicken pp60 c-src
42	26	1.3	1759	21	AAZ29700	Wild-type chicken
43	26	1.3	1759	22	AAH28357	Nucleotide sequenc
c 44	25	1.2	32	22	AAH41501	Human tyrosine kin
45	25	1.2	51	23	ABL00375	Human silent nonco

ALIGNMENTS

RESULT 1
ABK83939
ID ABK83939 standard; cDNA; 2015 BP.
XX
AC ABK83939;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #510.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 510; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2015 BP; 512 A; 540 C; 580 G; 383 T; 0 other;

Query Match 100.0%; Score 2015; DB 24; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080

Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320

Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980

Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
                ||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
```

RESULT 2
ABL66673
ID ABL66673 standard; DNA; 2015 BP.
XX
AC ABL66673;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:5010.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 5010; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 2015 BP; 512 A; 540 C; 580 G; 383 T; 0 other;

Query Match 100.0%; Score 2015; DB 24; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
```

```
Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320

Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980

Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
              ||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
```

```
RESULT 3
ABK83940
ID   ABK83940 standard; cDNA; 1926 BP.
XX
AC   ABK83940;
XX
DT   14-AUG-2002  (first entry)
XX
```

DE Human cDNA differentially expressed in granulocytic cells #511.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 511; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1926 BP; 497 A; 522 C; 520 G; 387 T; 0 other;

Query Match 77.0%; Score 1552; DB 24; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        85 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 144

Qy       238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       145 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 204

Qy       298 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       205 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 264

Qy       358 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       265 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 324

Qy       418 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       325 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 384

Qy       478 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       385 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 444

Qy       538 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 597
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       445 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 504

Qy       598 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 657
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       505 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 564

Qy       658 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 717
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       565 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 624

Qy       718 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 777
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       625 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 684

Qy       778 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 837
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       685 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 744

Qy       838 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 897
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       745 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 804

Qy       898 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 957
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       805 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 864

Qy       958 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1017
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       865 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 924

Qy      1018 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1077
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       925 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 984

Qy      1078 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1137
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       985 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1044

Qy      1138 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1197
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1104

Qy      1198 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1257
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1164

Qy       1258 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1317
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1224

Qy       1318 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1377
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1225 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1284

Qy       1378 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1437
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1344

Qy       1438 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1497
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1345 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1404

Qy       1498 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1557
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1405 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1464

Qy       1558 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1617
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1465 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1524

Qy       1618 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1677
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1525 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1584

Qy       1678 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1729
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1585 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1636
```

```
RESULT 4
ABL61214
ID   ABL61214 standard; DNA; 183 BP.
XX
AC   ABL61214;
XX
DT   04-SEP-2002  (first entry)
XX
DE   Human nucleotide fragment capable of inactivating HIV Nef protein.
XX
KW   Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW   pathogenicity; diagnosis; AIDS; human; ds.
XX
OS   Homo sapiens.
XX
PN   DE10109532-C1.
XX
PD   13-JUN-2002.
XX
PF   28-FEB-2001; 2001DE-1009532.
XX
PR   28-FEB-2001; 2001DE-1009532.
XX
PA   (GEYE/) GEYER M.
PA   (FACK/) FACKLER O.
XX
PI   Geyer M;
XX
DR   WPI; 2002-418264/45.
XX
PT   New fusion protein that blocks Nef protein, useful for treatment or
PT   diagnosis of acquired immune deficiency syndrome, has high specificity
PT   and affinity  -
XX
PS   Claim 12; Page 14; 22pp; German.
XX
CC   This invention describes a novel fusion protein for blocking the Nef
CC   protein of human immune deficiency virus (HIV) which comprises: (i)
CC   protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC   protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC   linker between protein domains 1 and 2. The products of the invention
CC   have virucide and anti-HIV activity and are capable of neutralising Nef,
CC   an accessory protein essential for pathogenicity of HIV-1. The fusion
CC   protein of the invention comprises the LL domain of the beta-subunit of
CC   the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC   tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC   of the invention are used for in vitro diagnosis of AIDS and for
CC   treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC   unlike HIV protease, has no human homologue, so the fusion protein (which
CC   binds Nef with very high affinity) should cause essentially no side
CC   effects. This sequence represents a human derived nucleotide fragment
CC   used in the construction of the fusion protein of the invention and which
CC   contains a PXXP-motif binding motif useful to the invention.
XX
SQ   Sequence 183 BP; 41 A; 50 C; 56 G; 36 T; 0 other;

  Query Match          9.1%;  Score 183;  DB 24;  Length 183;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-79;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        337 TCTGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTC 60

Qy        397 AGCTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCT 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCT 120

Qy        457 CGATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGAC 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGAC 180

Qy        517 TCT 519
              |||
Db        181 TCT 183
```

```
RESULT 5
ABL61215
ID   ABL61215 standard; DNA; 1416 BP.
XX
AC   ABL61215;
XX
DT   04-SEP-2002  (first entry)
XX
DE   Rat/human fusion construct capable of inactivating HIV Nef protein.
XX
KW   Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW   pathogenicity; diagnosis; AIDS; rat; human; ds.
XX
OS   Rattus sp.
OS   Homo sapiens.
OS   Synthetic.
XX
PN   DE10109532-C1.
XX
PD   13-JUN-2002.
XX
PF   28-FEB-2001; 2001DE-1009532.
XX
PR   28-FEB-2001; 2001DE-1009532.
XX
PA   (GEYE/) GEYER M.
PA   (FACK/) FACKLER O.
XX
PI   Geyer M;
XX
DR   WPI; 2002-418264/45.
XX
PT   New fusion protein that blocks Nef protein, useful for treatment or
```

```
PT   diagnosis of acquired immune deficiency syndrome, has high specificity
PT   and affinity  -
XX
PS   Claim 13; Page 14-15; 22pp; German.
XX
CC   This invention describes a novel fusion protein for blocking the Nef
CC   protein of human immune deficiency virus (HIV) which comprises: (i)
CC   protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC   protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC   linker between protein domains 1 and 2. The products of the invention
CC   have virucide and anti-HIV activity and are capable of neutralising Nef,
CC   an accessory protein essential for pathogenicity of HIV-1. The fusion
CC   protein of the invention comprises the LL domain of the beta-subunit of
CC   the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC   tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC   of the invention are used for in vitro diagnosis of AIDS and for
CC   treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC   unlike HIV protease, has no human homologue, so the fusion protein (which
CC   binds Nef with very high affinity) should cause essentially no side
CC   effects. This sequence represents a fusion construct composed of a rat
CC   nucleotide fragment which contains a dileucine (LL) motif and a human
CC   nucleotide fragment containing a PXXP-motif binding domain useful to the
CC   invention.
XX
SQ   Sequence 1416 BP; 340 A; 383 C; 386 G; 307 T; 0 other;

  Query Match             9.0%;  Score 181;  DB 24;  Length 1416;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-78;
  Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1233 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 1292

Qy        399 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1293 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 1352

Qy        459 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 1412

Qy        519 T 519
              |
Db       1413 T 1413


RESULT 6
ABL61216
ID   ABL61216 standard; DNA; 1542 BP.
XX
AC   ABL61216;
XX
DT   04-SEP-2002  (first entry)
XX
DE   Rat/human fusion construct capable of inactivating HIV Nef protein.
XX
KW   Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW   pathogenicity; diagnosis; AIDS; rat; human; ds.
XX
OS   Rattus sp.
OS   Homo sapiens.
OS   Synthetic.
XX
PN   DE10109532-C1.
XX
PD   13-JUN-2002.
XX
PF   28-FEB-2001; 2001DE-1009532.
XX
PR   28-FEB-2001; 2001DE-1009532.
XX
PA   (GEYE/) GEYER M.
PA   (FACK/) FACKLER O.
XX
PI   Geyer M;
XX
DR   WPI; 2002-418264/45.
XX
PT   New fusion protein that blocks Nef protein, useful for treatment or
PT   diagnosis of acquired immune deficiency syndrome, has high specificity
PT   and affinity  -
XX
PS   Claim 16; Page 15-16; 22pp; German.
XX
CC   This invention describes a novel fusion protein for blocking the Nef
CC   protein of human immune deficiency virus (HIV) which comprises: (i)
CC   protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC   protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC   linker between protein domains 1 and 2. The products of the invention
CC   have virucide and anti-HIV activity and are capable of neutralising Nef,
CC   an accessory protein essential for pathogenicity of HIV-1. The fusion
CC   protein of the invention comprises the LL domain of the beta-subunit of
CC   the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC   tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC   of the invention are used for in vitro diagnosis of AIDS and for
CC   treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC   unlike HIV protease, has no human homologue, so the fusion protein (which
CC   binds Nef with very high affinity) should cause essentially no side
CC   effects. This sequence represents a fusion construct composed of a rat
CC   nucleotide fragment which contains a dileucine (LL) motif and a human
CC   nucleotide fragment containing a PXXP-motif binding domain useful to the
CC   invention.
XX
SQ   Sequence 1542 BP; 369 A; 419 C; 427 G; 327 T; 0 other;

  Query Match             9.0%;  Score 181;  DB 24;  Length 1542;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-78;
  Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1290 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 1349

Qy        399 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 1409

Qy        459 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1410 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 1469

Qy        519 T 519
              |
Db       1470 T 1470


RESULT 7
AAT19957
ID   AAT19957 standard; cDNA to mRNA; 369 BP.
XX
AC   AAT19957;
XX
DT   17-JUL-1996  (first entry)
XX
DE   Human gene signature HUMGS01089.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
```

```
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 520; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 369 BP; 82 A; 97 C; 102 G; 75 T; 13 other;

  Query Match             8.4%;  Score 169;  DB 16;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-72;
  Matches  169;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1535 CTCGCCCAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACC 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 CTCGCCCAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACC 92

Qy       1595 GTCCGGAGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGG 1654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GTCCGGAGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGG 152

Qy       1655 CCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCA 1703
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 CCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCA 201


RESULT 8
ABA50558
ID   ABA50558 standard; DNA; 171 BP.
XX
AC   ABA50558;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9253.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 4; SEQ ID NO 9253; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 9
ABA68516
ID   ABA68516 standard; DNA; 171 BP.
XX
AC   ABA68516;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16821.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 16821; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 10
ABA35497
ID   ABA35497 standard; DNA; 171 BP.
XX
AC   ABA35497;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #13963 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 13963; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 11
AAK16884
ID   AAK16884 standard; DNA; 171 BP.
XX
AC   AAK16884;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16875.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
```

```
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 16875; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 12
AAK42654
ID   AAK42654 standard; DNA; 171 BP.
XX
AC   AAK42654;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17211.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 17211; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 13
AAI23408
ID   AAI23408 standard; DNA; 171 BP.
XX
AC   AAI23408;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13341 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
```

```
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 13341; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 14
AAI48728
ID   AAI48728 standard; DNA; 171 BP.
XX
AC   AAI48728;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17414 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 17414; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 15
AAI09035
ID   AAI09035 standard; DNA; 171 BP.
XX
AC   AAI09035;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #9026 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 9026; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
```

CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

Query Match 6.6%; Score 133; DB 22; Length 171;
Best Local Similarity 100.0%; Pred. No. 7.1e-55;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133
```

Search completed: July 4, 2003, 02:31:01
Job time : 458 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 00:38:48 ; Search time 5238 Seconds
(without alignments)
11195.524 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2015	100.0	2015	6	AX334501	AX334501 Sequence
2		2015	100.0	2015	9	HUMHCKA	M16591 Human hemop
3		1606	79.7	2044	9	BC014435	BC014435 Homo sapi
4		1555	77.2	2105	9	AK026432	AK026432 Homo sapi
5		1552	77.0	1926	9	HUMHCKB	M16592 Human hemop
6		423	21.0	4507	9	HSHCKE12	X58743 H.sapiens H
7		303	15.0	111694	9	HSJ836N17	AL049539 Human DNA
8		274	13.6	333	11	G06122	G06122 human STS W
9		182	9.0	5268	9	HSHCKE69	X58741 H.sapiens H
10		157	7.8	2167	9	HSHCKE11	X58742 H.sapiens H
11		112	5.6	25010	9	AL353092	AL353092 Human DNA
12		107	5.3	1515	9	MFA320181	AJ320181 Macaca fa
13		82	4.1	366	11	G25924	G25924 human STS E
14		77	3.8	958	9	HUMHCK	M73233 Human hemop
15		76	3.8	10348	9	AB071605	AB071605 Homo sapi
16		68	3.4	1911	6	AX401935	AX401935 Sequence
17		68	3.4	1911	10	RATHCTK	M83666 Rattus norv
18		68	3.4	1911	10	S74141	S74141 hck=tyrosin
19		68	3.4	1940	10	RRHCKMR	X62345 R.rattus hc
20	c	65	3.2	1926	9	HUMHCKB	M16592 Human hemop
21	c	56	2.8	86196	9	AL592046	AL592046 Human DNA
22	c	56	2.8	169913	2	AC031980	AC031980 Homo sapi
23		47	2.3	1960	10	MMHCK	Y00487 Mouse hck g
24		47	2.3	2002	10	MUSBMK	J03023 Murine macr
25		47	2.3	2104	10	BC010478	BC010478 Mus muscu
26		47	2.3	200329	10	AC078911	AC078911 Mus muscu
27		47	2.3	208718	2	AL807380	AL807380 Mus muscu
28		32	1.6	2298	5	AF321110	AF321110 Salmo sal
29	c	32	1.6	211607	2	AC094844	AC094844 Rattus no
30		30	1.5	145542	2	AC087618	AC087618 Homo sapi
31		30	1.5	191622	2	AC022239	AC022239 Homo sapi
32	c	29	1.4	301	11	HS180I13S	AL031189 H.sapiens
33		29	1.4	1965	9	HSM804186	AL832875 Homo sapi
34		29	1.4	2210	9	BC007371	BC007371 Homo sapi
35		29	1.4	2235	9	HSBITPTK	Z33998 H.sapiens m
36		29	1.4	2251	9	BC032413	BC032413 Homo sapi
37		29	1.4	2608	9	S76617	S76617 blk=protein
38		29	1.4	211607	2	AC094844	AC094844 Rattus no
39		28	1.4	1518	9	HUMLYNTK	M79321 Human Lyn B
40		28	1.4	2298	9	HUMLYN	M16038 Human lyn m
41	c	28	1.4	184349	9	AC046176	AC046176 Homo sapi
42		27	1.3	180424	2	AC090496	AC090496 Mus muscu
43		26	1.3	72	14	ALRSRCB	J02351 Rous sarcom
44		26	1.3	1016	14	REASV3	V01167 Avian sarco
45		26	1.3	1578	14	S37068	S37068 src (tsUP1)

ALIGNMENTS

RESULT 1
AX334501
LOCUS AX334501 2015 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 5010 from Patent WO0194629.
ACCESSION AX334501
VERSION AX334501.1 GI:18125220
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature

```
            gene sets
  JOURNAL   Patent: WO 0194629-A 5010 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES              Location/Qualifiers
     source           1. .2015
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT     512 a    540 c    580 g    383 t
ORIGIN
```

Query Match 100.0%; Score 2015; DB 6; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080

Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320

Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
```

```
Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
             ||||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015


RESULT 2
HUMHCKA
LOCUS       HUMHCKA                 2015 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human hemopoietic cell protein-tyrosine kinase (HCK) gene, complete
            cds, clone lambda-a2/1a.
ACCESSION   M16591
VERSION     M16591.1  GI:183911
KEYWORDS    kinase; protein kinase; protein-tyrosine kinase.
SOURCE      Human hemopoietic cell, cDNA to mRNA, clone lambda-a2/1a.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2015)
  AUTHORS   Quintrell,N., Lebo,R., Varmus,H., Bishop,J.M., Pettenati,M.J., Le
            Beau,M.M., Diaz,M.O. and Rowley,J.D.
  TITLE     Identification of a human gene (HCK) that encodes a
            protein-tyrosine kinase and is expressed in hemopoietic cells
  JOURNAL   Mol. Cell. Biol. 7 (6), 2267-2275 (1987)
  MEDLINE   87257942
   PUBMED   3496523
FEATURES             Location/Qualifiers
     source          1. .2015
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="20q11-q12"
     gene            1. .2015
                     /gene="HCK"
     mRNA            <1. .2015
                     /gene="HCK"
                     /product="HCK mRNA"
     CDS             169. .1686
                     /gene="HCK"
                     /note="protein-tyrosine kinase"
                     /codon_start=1
                     /protein_id="AAA52643.1"
                     /db_xref="GI:306832"
                     /db_xref="GDB:G00-119-303"
                     /translation="MGSMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNS
                     HNSNTPGIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATR
                     KEGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYS
                     LSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSV
                     PCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSM
                     SVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
                     PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTA
                     REGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALE
                     RGYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
BASE COUNT      512 a    540 c    580 g    383 t
ORIGIN      130 bp upstream of BamHI site; chromosome 20q11-q12.

  Query Match           100.0%;  Score 2015;  DB 9;  Length 2015;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2015;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080

Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
```

```
Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980

Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
             ||||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
```

```
RESULT 3
BC014435
LOCUS       BC014435               2044 bp    mRNA    linear   PRI 19-SEP-2001
DEFINITION  Homo sapiens, clone MGC:22922 IMAGE:4855747, mRNA, complete cds.
ACCESSION   BC014435
VERSION     BC014435.1  GI:15680176
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2044)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 34 Row: d Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 10439295.
FEATURES             Location/Qualifiers
     source          1. .2044
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:22922 IMAGE:4855747"
                     /tissue_type="Primary B-Cells from Tonsils"
                     /clone_lib="NIH_MGC_48"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             187. .1704
                     /codon_start=1
                     /product="Unknown (protein for MGC:22922)"
                     /protein_id="AAH14435.1"
                     /db_xref="GI:15680177"
                     /translation="MGCMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNS
                     HNSNTPGIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATR
                     KEGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYS
                     LSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSV
                     PCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSM
                     SVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
                     PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTA
                     REGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALE
                     RGYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
BASE COUNT      523 a    572 c    566 g    383 t
ORIGIN
```

```
  Query Match             79.7%;  Score 1606;  DB 9;  Length 2044;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1826;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 255

Qy        238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 315

Qy        298 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 375

Qy        358 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 435

Qy        418 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 495

Qy        478 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 555

Qy        538 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        556 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 615

Qy        598 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 675

Qy        658 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 735

Qy        718 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 795

Qy        778 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 855

Qy        838 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 915

Qy        898 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 975

Qy        958 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1035

Qy       1018 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1095

Qy       1078 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1155

Qy       1138 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1215

Qy       1198 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1216 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1275

Qy       1258 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1335

Qy       1318 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1395

Qy       1378 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1455

Qy       1438 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1515

Qy       1498 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1575

Qy       1558 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1635

Qy       1618 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1695

Qy       1678 CAGCCATGATAGGGAGGACCAGGGCAGGG-CAGGGGGTGCCCAGGTGGTGGCTCGAAGGT 1736
              |||||||||||||||||||||||||||| |||||||||||||||||||||||  |||||
Db       1696 CAGCCATGATAGGGAGGACCAGGGCAGGGCCAGGGGGTGCCCAGGTGGTGGCTGCAAGGT 1755

Qy       1737 GGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCG 1796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 GGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCG 1815

Qy       1797 CTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTG 1856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 CTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTG 1875

Qy       1857 ACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCC 1916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 ACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCC 1935

Qy       1917 TGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAA 1976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 TGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAA 1995

Qy       1977 ATGAATATTTAAATAAAAGATATAAATGC 2005
              |||||||||||||||||||||||||||||
Db       1996 ATGAATATTTAAATAAAAGATATAAATGC 2024
```

```
RESULT 4
AK026432
LOCUS       AK026432               2105 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ22779 fis, clone KAIA1741.
ACCESSION   AK026432
VERSION     AK026432.1  GI:10439295
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens ileal mucosa cDNA to mRNA, clone_lib:kaia
            clone:KAIA1741.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
            Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
            Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
            Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
            Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2105)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2105
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KAIA1741"
                     /tissue_type="ileal mucosa"
                     /clone_lib="kaia"
                     /note="cloning vector pME18SFL3"
     CDS             235. .1752
                     /note="unnamed protein product"
                     /codon_start=1
```

```
                     /protein_id="BAB15482.1"
                     /db_xref="GI:10439296"
                     /translation="MGCMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNS
                     HNSNTPGIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATR
                     KEGYIPSNYVARVDSLETEERFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYS
                     LSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSV
                     PCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSM
                     SVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
                     PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTA
                     REGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALE
                     RGYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
BASE COUNT      541 a    585 c    587 g    392 t
ORIGIN

  Query Match             77.2%;  Score 1555;  DB 9;  Length 2105;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1825;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 303

Qy        238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 363

Qy        298 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 423

Qy        358 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 483

Qy        418 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 543

Qy        478 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 537
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        544 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGAGG 603

Qy        538 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 663

Qy        598 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 723

Qy        658 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 783

Qy        718 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 843

Qy        778 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 903

Qy        838 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 963

Qy        898 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 1023

Qy        958 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1083

Qy       1018 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1143

Qy       1078 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1203

Qy       1138 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1263

Qy       1198 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1323

Qy       1258 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1383

Qy       1318 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1443

Qy       1378 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1503

Qy       1438 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1563

Qy       1498 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1564 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1623

Qy       1558 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1624 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1683

Qy       1618 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1684 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1743

Qy       1678 CAGCCATGATAGGGAGGACCAGGGCAGGG-CAGGGGGTGCCCAGGTGGTGGCTCGAAGGT 1736
              |||||||||||||||||||||||||||| |||||||||||||||||||| |||||
Db       1744 CAGCCATGATAGGGAGGACCAGGGCAGGGCCAGGGGGTGCCCAGGTGGTGGCTGCAAGGT 1803

Qy       1737 GGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCG 1796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 GGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCG 1863

Qy       1797 CTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTG 1856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 CTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTG 1923

Qy       1857 ACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCC 1916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 ACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCC 1983

Qy       1917 TGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAA 1976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 TGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAA 2043

Qy       1977 ATGAATATTTAAATAAAAGATATAAATGC 2005
              |||||||||||||||||||||||||||||
Db       2044 ATGAATATTTAAATAAAAGATATAAATGC 2072


RESULT 5
HUMHCKB
LOCUS       HUMHCKB                 1926 bp    mRNA    linear   PRI 08-NOV-1994
```

```
DEFINITION  Human hemopoietic cell protein-tyrosine kinase (HCK) gene, complete
            cds, clone HK24.
ACCESSION   M16592
VERSION     M16592.1  GI:183913
KEYWORDS    kinase; protein kinase; protein-tyrosine kinase.
SOURCE      Human mitogen-stimulated leukocyte, cDNA to mRNA, clone HK24.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1926)
  AUTHORS   Ziegler,S.F., Marth,J.D., Lewis,D.B. and Perlmutter,R.M.
  TITLE     Novel protein-tyrosine kinase gene (hck) preferentially expressed
            in cells of hematopoietic origin
  JOURNAL   Mol. Cell. Biol. 7 (6), 2276-2285 (1987)
  MEDLINE   87257943
   PUBMED   3453117
FEATURES             Location/Qualifiers
     source          1. .1926
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="20q11-q12"
     gene            1. .1926
                     /gene="HCK"
     mRNA            <1. .1926
                     /gene="HCK"
                     /product="HCK mRNA"
     CDS             76. .1593
                     /gene="HCK"
                     /note="protein-tyrosine kinase"
                     /codon_start=1
                     /protein_id="AAA52644.1"
                     /db_xref="GI:306833"
                     /db_xref="GDB:G00-119-303"
                     /translation="MGCMKSKFLQVGGNTFSKTETSASPHCPVYVPDPTSTIKPGPNS
                     HNSNTPGIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATR
                     KEGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYS
                     LSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSV
                     PCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSM
                     SVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPL
                     PKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTA
                     REGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALE
                     RGYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
BASE COUNT      497 a    522 c    520 g    387 t
ORIGIN      1 bp upstream of EcoRI site; chromosome 20q11-q12.

  Query Match             77.0%;  Score 1552;  DB 9;  Length 1926;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1552;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 144

Qy        238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 204

Qy        298 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 264

Qy        358 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 324

Qy        418 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 384

Qy        478 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 444

Qy        538 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 504

Qy        598 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 564

Qy        658 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 624

Qy        718 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 684

Qy        778 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 744

Qy        838 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 804

Qy        898 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 864

Qy        958 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 924

Qy       1018 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 984

Qy       1078 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1044

Qy       1138 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1104

Qy       1198 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1164

Qy       1258 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1224

Qy       1318 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1225 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1284

Qy       1378 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1344

Qy       1438 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1345 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1404

Qy       1498 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1405 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1464

Qy       1558 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1465 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1524

Qy       1618 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1677
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1525 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1584

Qy       1678 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1729
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1585 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1636


RESULT 6
HSHCKE12
LOCUS       HSHCKE12                4507 bp    DNA     linear   PRI 27-AUG-1999
DEFINITION  H.sapiens HCK gene for tyrosine kinase (PTK), exon 12.
ACCESSION   X58743 X59743
VERSION     X58743.1  GI:32044
KEYWORDS    proto-oncogene; src family; T-cell receptor alpha-chain; Tyrosine
            kinase; V-alpha gene segment; variable region.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4507)
  AUTHORS   Hradetzky,D., Strebhardt,K. and Rubsamen-Waigmann,H.
  TITLE     The genomic locus of the human hemopoietic-specific cell protein
            tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of
            exon-intron structure among human PTKs of the src family
  JOURNAL   Gene 113 (2), 275-280 (1992)
  MEDLINE   92241680
   PUBMED   1572549
REFERENCE   2  (bases 1 to 4507)
  AUTHORS   Hradetzky,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1991) D. Hradetzky, Chemotherapeutisches
            Forschungsinstitut, Georg-Speyer-Haus, Paul Ehrlich Str 42-44, 6000
            Frankfurt 70, Federal Republic of Germany
COMMENT     See also X58736-X58740, X58744-X58769
            See also X58741 and X58742.
FEATURES             Location/Qualifiers
     source          1. .4507
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q11-12"
                     /clone="6500 H"
                     /tissue_type="spleen"
                     /clone_lib="genomic; TS48"
                     /dev_stage="adult"
     intron          <1. .3806
                     /number=11
     repeat_region   1982. .2149
                     /note="ALU repeat partial"
     repeat_region   2390. .2748
                     /note="ALU repeat VIII"
     gene            3807. .4342
                     /gene="HCK"
     exon            3807. .4342
                     /gene="HCK"
                     /number=12
                     /usedin=X58741:HCK_cds
                     /usedin=X58741:HCK_mRNA
     polyA_signal    4311. .4317
                     /gene="HCK"
BASE COUNT     1159 a   1064 c   1095 g   1189 t
ORIGIN

  Query Match           21.0%;  Score 423;  DB 9;  Length 4507;
  Best Local Similarity   99.6%;  Pred. No. 3.6e-232;
  Matches  523;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1481 CAGGGATGTCAAACCCTGAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCC 1540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3804 CAGGGATGTCAAACCCTGAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCC 3863

Qy       1541 CAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGG 1600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3864 CAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGG 3923

Qy       1601 AGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAG 1660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3924 AGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAG 3983

Qy       1661 AGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAG 1720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3984 AGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAG 4043

Qy       1721 GTGGTGGCTCGAAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCC 1780
                |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4044 GTGGTGGCTGCAAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCC 4103

Qy       1781 CAGACACCCACCCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACT 1840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4104 CAGACACCCACCCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACT 4163

Qy       1841 GGAAAATCTCTTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGT 1900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4164 GGAAAATCTCTTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGT 4223

Qy       1901 TGATATTTCTATTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAA 1960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4224 TGATATTTCTATTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAA 4283

Qy       1961 CTTTCAAAATAGTGAAATGAATATTTAAATAAAAGATATAAATGC 2005
                |||||||||||||||||||||||||||||||||||||||||||
Db       4284 CTTTCAAAATAGTGAAATGAATATTTAAATAAAAGATATAAATGC 4328


RESULT 7
HSJ836N17
LOCUS       HSJ836N17               111694 bp    DNA     linear   PRI 01-MAR-2001
DEFINITION  Human DNA sequence from clone RP5-836N17 on chromosome
            20q11.1-11.21 Contains part of the HCK (hemopoietic cell kinase)
            gene, the KIAA0255 gene, a ribosomal protein L30 pseudogene, ESTs,
            STSs, GSSs and CpG Islands, complete sequence.
ACCESSION   AL049539
VERSION     AL049539.21  GI:6456855
KEYWORDS    HTG; CpG island; HCK; hemopoietic; KIAA0255; kinase; RPL30.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 111694)
  AUTHORS   Sehra,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 21, 1999 this sequence version replaced gi:6433925.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence is
            the entire insert of clone RP5-836N17 The true left end of clone
            RP11-392M18 is at 77067 in this sequence. This sequence was
            generated from part of bacterial clone contigs of human chromosome
            20, constructed by the Sanger Centre Chromosome 20 Mapping Group.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
```

```
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by at least
          one plasmid subclone or more than one M13 subclone; and the
          assembly was confirmed by restriction digest. RP5-836N17 is from
          the library RPCI-5 constructed by the group of Pieter de Jong. For
          further details see
          http://www.chori.org/bacpac/home.htm
          VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .111694
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q11.1-11.21"
                     /clone="RP5-836N17"
                     /clone_lib="RPCI-5"
     repeat_region   1. .62
                     /note="AluJ/FLAM repeat: matches 1. .62 of consensus"
     repeat_region   73. .173
                     /note="L2 repeat: matches 2597. .2709 of consensus"
     repeat_region   243. .302
                     /note="15 copies 4 mer catc 98% conserved"
     repeat_region   601. .898
                     /note="AluSq repeat: matches 1. .295 of consensus"
     repeat_region   1188. .1345
                     /note="MER5A repeat: matches 1. .185 of consensus"
     repeat_region   1364. .1668
                     /note="AluSx repeat: matches 1. .307 of consensus"
     repeat_region   2074. .2381
                     /note="AluSp repeat: matches 3. .313 of consensus"
     repeat_region   2583. .2904
                     /note="AluY repeat: matches 1. .311 of consensus"
     repeat_region   2993. .3134
                     /note="MIR repeat: matches 2. .143 of consensus"
     repeat_region   3135. .3439
                     /note="AluSx repeat: matches 1. .305 of consensus"
     repeat_region   3440. .3503
                     /note="MIR repeat: matches 143. .206 of consensus"
     repeat_region   3491. .3594
                     /note="MIR repeat: matches 60. .164 of consensus"
     repeat_region   3866. .4024
                     /note="MER69 repeat: matches 66. .236 of consensus"
     repeat_region   4033. .4126
                     /note="MER69 repeat: matches 2422. .2510 of consensus"
     misc_feature    complement(4261. .4710)
                     /note="match: GSS: Em:AQ339627"
     gene            4334. .26412
                     /gene="HCK"
     mRNA            join(<4334. .4437,8454. .8603,8951. .9103,11188. .11367,
                     13128. .13204,18423. .18576,23564. .23695,25877. .26412)
                     /gene="HCK"
                     /product="dJ836N17.1 (hemopoietic cell kinase)"
                     /note="match: cDNAs: Em:J03023 Em:M83666 Em:S74141
                     Em:X62345 Em:Y00487 Em:M12056 Em:M57697 Em:M36881
                     Em:M57696 Em:U23852 Em:M30903 Em:U01349 Em:X03533
                     Em:L14823 Em:X52191 Em:U70324 Em:M19722 Em:M17031
                     Em:X15345 Em:U35365 Em:M79321 Em:J03579 Em:AF000302
                     Em:M64608 Em:AF000301 Em:X13529 Em:AF000300 Em:X12461
                     Em:M23422 Em:L14951 Em:X54970 Em:X54971 Em:M14676
                     Em:M27266 Em:AF130457 Em:X52822 Em:S76617 Em:X16440
                     Em:M16592 Em:M16591 Em:M16038 Em:M14333 Em:L14782
                     Em:AF081803 Em:X05027 Em:X57018 Em:X13207 Em:Z33998
                     Em:U07236 Em:X67786 Em:X67677 Em:M24704
                     match: ESTs: Em:AA763708 Em:AA149096 Em:W87315 Em:AI912730
                     Em:AI220607 Em:AI572095"
                     /evidence=not_experimental
     CDS             join(<4334. .4437,8454. .8603,8951. .9103,11188. .11367,
                     13128. .13204,18423. .18576,23564. .23695,25877. .26079)
                     /gene="HCK"
                     /note="Continues in Em:AL353092 as dJ180I13.1
                     match: proteins: Sw:P06239 Sw:P06240 Sw:P08103 Sw:P50545
                     Tr:O13064 Sw:P42683 Sw:P08631 Sw:Q07014 Sw:P07948"
                     /codon_start=2
                     /evidence=not_experimental
                     /product="dJ836N17.1 (hemopoietic cell kinase)"
                     /protein_id="CAB75606.1"
                     /db_xref="GI:7018398"
                     /db_xref="SPTREMBL:Q9NUA4"
                     /translation="WFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYSLSVRDY
                     DPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKLSVPCMSSK
                     PQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSMSVEAFL
                     AEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPLPKLIDF
                     SAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARVIEDNEYTAREGAKF
                     PIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALERGYRMP
                     RPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
     repeat_region   4751. .5056
                     /note="AluSc repeat: matches 1. .306 of consensus"
     repeat_region   5059. .5348
                     /note="AluJb repeat: matches 24. .295 of consensus"
     repeat_region   5406. .5666
                     /note="AluJo repeat: matches 41. .300 of consensus"
     misc_feature    complement(5985. .6147)
                     /note="match: STS: Em:AA443921"
     repeat_region   6602. .6796
                     /note="MIR repeat: matches 10. .205 of consensus"
     repeat_region   6860. .6919
                     /note="MIR repeat: matches 187. .242 of consensus"
     repeat_region   6920. .7219
                     /note="AluSx repeat: matches 1. .300 of consensus"
     repeat_region   7220. .7328
                     /note="MIR repeat: matches 64. .187 of consensus"
     repeat_region   7378. .7456
                     /note="MADE1 repeat: matches 2. .80 of consensus"
     repeat_region   7461. .7926
                     /note="MLT1D repeat: matches 1. .505 of consensus"
     repeat_region   8095. .8311
                     /note="MIR repeat: matches 35. .261 of consensus"
     repeat_region   9492. .9524
                     /note="L2 repeat: matches 2652. .2686 of consensus"
     repeat_region   9791. .10099
                     /note="AluSx repeat: matches 1. .308 of consensus"
     repeat_region   10223. .10555
                     /note="L2 repeat: matches 2141. .2482 of consensus"
     repeat_region   10576. .10867
                     /note="AluY repeat: matches 12. .303 of consensus"
     repeat_region   10880. .11054
                     /note="MIR repeat: matches 21. .188 of consensus"
     repeat_region   11862. .12205
                     /note="AluJo repeat: matches 257. .302 of consensus"
     repeat_region   12206. .12501
                     /note="AluY repeat: matches 1. .295 of consensus"
     repeat_region   12502. .12764
                     /note="AluJo repeat: matches 1. .257 of consensus"
     repeat_region   13048. .13092
                     /note="MADE1 repeat: matches 36. .79 of consensus"
     misc_feature    complement(13162. .13462)
                     /note="match: STS: Em:HS180I13S"
     misc_feature    complement(13162. .13412)
                     /note="match: STS: Em:AL031189"
     repeat_region   13302. .13417
                     /note="MIR repeat: matches 151. .262 of consensus"
     repeat_region   13418. .13810
                     /note="MLT1H repeat: matches 99. .540 of consensus"
     repeat_region   13835. .13922
                     /note="MIR repeat: matches 101. .182 of consensus"
     repeat_region   13921. .13981
                     /note="MIR repeat: matches 76. .137 of consensus"
     repeat_region   13951. .14131
                     /note="L2 repeat: matches 2505. .2726 of consensus"
     repeat_region   14285. .14527
                     /note="L2 repeat: matches 2300. .2551 of consensus"
     repeat_region   15366. .15471
                     /note="MIR repeat: matches 23. .135 of consensus"
     repeat_region   15483. .15707
                     /note="MIR repeat: matches 1. .256 of consensus"
     repeat_region   15855. .16037
```

```
                     /note="MIR repeat: matches 12. .193 of consensus"
     repeat_region   16268. .16567
                     /note="AluJo repeat: matches 5. .312 of consensus"
     repeat_region   16872. .16920
                     /note="MIR repeat: matches 125. .173 of consensus"
     repeat_region   17077. .17497
                     /note="L2 repeat: matches 2098. .2595 of consensus"
     repeat_region   17498. .17804
                     /note="AluSx repeat: matches 1. .307 of consensus"
     misc_feature    complement(17780. .18196)
                     /note="match: GSS: Em:AQ590401"
     repeat_region   17805. .17925
                     /note="L2 repeat: matches 2595. .2710 of consensus"
     misc_feature    complement(17926. .18196)
                     /note="match: GSS: Em:AQ590401"

  Query Match             15.0%;  Score 303;  DB 9;  Length 111694;
  Best Local Similarity   99.4%;  Pred. No. 1e-162;
  Matches  523;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1481 CAGGGATGTCAAACCCTGAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCC 1540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25874 CAGGGATGTCAAACCCTGAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCC 25933

Qy       1541 CAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGG 1600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25934 CAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGG 25993

Qy       1601 AGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAG 1660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25994 AGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAG 26053

Qy       1661 AGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGG-CAGGGGGTGCCCA 1719
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      26054 AGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCAGGGCCAGGGGGTGCCCA 26113

Qy       1720 GGTGGTGGCTCGAAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTC 1779
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26114 GGTGGTGGCTGCAAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTC 26173

Qy       1780 CCAGACACCCACCCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGAC 1839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26174 CCAGACACCCACCCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGAC 26233

Qy       1840 TGGAAAATCTCTTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAG 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26234 TGGAAAATCTCTTTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAG 26293

Qy       1900 TTGATATTTCTATTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAA 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26294 TTGATATTTCTATTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAA 26353

Qy       1960 ACTTTCAAAATAGTGAAATGAATATTTAAATAAAAGATATAAATGC 2005
             ||||||||||||||||||||||||||||||||||||||||||||||
Db      26354 ACTTTCAAAATAGTGAAATGAATATTTAAATAAAAGATATAAATGC 26399


RESULT 8
G06122
LOCUS       G06122                   333 bp    DNA     linear   STS 19-OCT-1995
DEFINITION  human STS WI-7020, sequence tagged site.
ACCESSION   G06122
VERSION     G06122.1  GI:859367
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens STSs derived from sequences in dbEST and the Unigene
            collection.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 333)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped ESTs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: TAGGGAGGACCAGGGCAG
            Primer B: TGGTAAAGACTTTGGCATTTATATC
            STS size: 333
            PCR Profile:
                 Presoak:
                 Denaturation:
                 Annealing: 56 degrees C
                 Polymerization:
                 PCR Cycles: 35
                 Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
                MgCl2: 1.5 mM
                KCl: 50 mM
                Tris-HCL: 10 mM
                pH: 9.3

            Prepared with primer pairs derived from X58743 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="734_E_?; 908_C_6; 768_C_11; (808,809)_(G,A)_(1,6)"
     STS             1. .333
     primer_bind     1. .18
     primer_bind     complement(309. .333)
BASE COUNT       87 a     83 c     77 g     86 t
ORIGIN

  Query Match             13.6%;  Score 274;  DB 11;  Length 333;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-146;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1732 AAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCAC 1791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 AAGGTGGCTCCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCAC 105

Qy       1792 CCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCT 1851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CCTCGCTTCAGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCT 165

Qy       1852 TTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTA 1911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TTTTGACTCTTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTA 225

Qy       1912 TTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATA 1971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 TTTCCTGGAATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATA 285

Qy       1972 GTGAAATGAATATTTAAATAAAAGATATAAATGC 2005
             ||||||||||||||||||||||||||||||||||
Db        286 GTGAAATGAATATTTAAATAAAAGATATAAATGC 319


RESULT 9
HSHCKE69
LOCUS       HSHCKE69                 5268 bp    DNA     linear   PRI 27-AUG-1999
```

```
DEFINITION  H.sapiens HCK gene for tyrosine kinase (PTK), exons 6-9.
ACCESSION   X58741 X59741
VERSION     X58741.1  GI:32045
KEYWORDS    proto-oncogene; src family; T-cell receptor alpha-chain; Tyrosine
            kinase; V-alpha gene segment; variable region.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5268)
  AUTHORS   Hradetzky,D., Strebhardt,K. and Rubsamen-Waigmann,H.
  TITLE     The genomic locus of the human hemopoietic-specific cell protein
            tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of
            exon-intron structure among human PTKs of the src family
  JOURNAL   Gene 113 (2), 275-280 (1992)
  MEDLINE   92241680
   PUBMED   1572549
REFERENCE   2  (bases 1 to 5268)
  AUTHORS   Hradetzky,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1991) D. Hradetzky, Chemotherapeutisches
            Forschungsinstitut, Georg-Speyer-Haus, Paul Ehrlich Str 42-44, 6000
            Frankfurt 70, Federal Republic of Germany
COMMENT     See also X58736-X58740, X58744-X58769
            See also X58742 and X58743.
FEATURES             Location/Qualifiers
     source          1. .5268
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q11-12"
                     /clone="H530R"
                     /tissue_type="spleen"
                     /clone_lib="genomic; TS48"
                     /dev_stage="adult"
     intron          <1. .835
                     /number=5
     gene            join(836. .5227,X58742.1:16. .1065,X58743.1:3807. .4342)
                     /gene="HCK"
     mRNA            join(836. .985,1333. .1485,3572. .3751,5151. .5227,
                     X58742.1:16. .169,X58742.1:934. .1065,X58743.1:3807. .4342)
                     /partial
                     /gene="HCK"
                     /label=HCK_mRNA
     CDS             join(836. .985,1333. .1485,3572. .3751,5151. .5227,
                     X58742.1:16. .169,X58742.1:934. .1065,X58743.1:3807. .4009)
                     /partial
                     /gene="HCK"
                     /codon_start=3
                     /label=HCK_cds
                     /product="tyrosine kinase"
                     /protein_id="CAA41565.2"
                     /db_xref="GI:5804911"
                     /translation="SYSLSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQEL
                     VDHYKKGNDGLCQKLSVPCMSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMAT
                     YNKHTKVAVKTMKPGSMSVEAFLAEANVMKTLQHDKLVKLHAVVTKEPIYIITEFMAK
                     GSLLDFLKSDEGSKQPLPKLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCK
                     IADFGLARVIEDNEYTAREGAKFPIKWTAPEAINFGSFTIKSDVWSFGILLMEIVTYG
                     RIPYPGMSNPEVIRALERGYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLD
                     DFYTATESQYQQQP"
     exon            836. .985
                     /gene="HCK"
                     /number=6
     intron          986. .1332
                     /gene="HCK"
                     /number=6
     exon            1333. .1485
                     /gene="HCK"
                     /number=7
     intron          1486. .3571
                     /gene="HCK"
                     /number=7
     repeat_region   2182. .2482
                     /note="ALU repeat I"
     repeat_region   2948. .3248
                     /note="ALU repeat II"
     exon            3572. .3751
                     /gene="HCK"
                     /number=8
     intron          3752. .5150
                     /gene="HCK"
                     /number=8
     repeat_region   4243. .4528
                     /note="ALU repeat III"
     repeat_region   4527. .4784
                     /note="ALU repeat IV"
     exon            5151. .5227
                     /gene="HCK"
                     /number=9
     intron          5228. .>5268
                     /number=9
BASE COUNT     1330 a   1265 c   1382 g   1291 t
ORIGIN

  Query Match             9.0%;  Score 182;  DB 9;  Length 5268;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-93;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        940 GCCACCTACAACAAGCACACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCG 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3571 GCCACCTACAACAAGCACACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCG 3630

Qy       1000 GTGGAGGCCTTCCTGGCAGAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTC 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3631 GTGGAGGCCTTCCTGGCAGAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTC 3690

Qy       1060 AAACTTCATGCGGTGGTCACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAA 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3691 AAACTTCATGCGGTGGTCACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAA 3750

Qy       1120 GG 1121
              ||
Db       3751 GG 3752


RESULT 10
HSHCKE11
LOCUS       HSHCKE11                2167 bp    DNA     linear   PRI 27-AUG-1999
DEFINITION  H.sapiens HCK gene for tyrosine kinase (PTK), exons 10-11.
ACCESSION   X58742 X59742
VERSION     X58742.1  GI:32043
KEYWORDS    proto-oncogene; src family; T-cell receptor alpha-chain; Tyrosine
            kinase; V-alpha gene segment; variable region.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2167)
  AUTHORS   Hradetzky,D., Strebhardt,K. and Rubsamen-Waigmann,H.
  TITLE     The genomic locus of the human hemopoietic-specific cell protein
            tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of
            exon-intron structure among human PTKs of the src family
  JOURNAL   Gene 113 (2), 275-280 (1992)
  MEDLINE   92241680
   PUBMED   1572549
REFERENCE   2  (bases 1 to 2167)
  AUTHORS   Hradetzky,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1991) D. Hradetzky, Chemotherapeutisches
            Forschungsinstitut, Georg-Speyer-Haus, Paul Ehrlich Str 42-44, 6000
            Frankfurt 70, Federal Republic of Germany
COMMENT     See also X58736-X58740, X58744-X58769
            See also X58741 and x58743.
FEATURES             Location/Qualifiers
     source          1. .2167
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q11-12"
                     /clone="D640 H"
                     /tissue_type="spleen"
                     /clone_lib="genomic; TS48"
                     /dev_stage="adult"
     intron          <1. .15
                     /number=9
     gene            16. .1065
                     /gene="HCK"
     exon            16. .169
                     /gene="HCK"
                     /number=10
                     /usedin=X58741:HCK_cds
                     /usedin=X58741:HCK_mRNA
     intron          170. .933
                     /gene="HCK"
                     /number=10
     repeat_region   203. .430
                     /note="ALU repeat V"
     exon            934. .1065
                     /gene="HCK"
                     /number=11
                     /usedin=X58741:HCK_cds
                     /usedin=X58741:HCK_mRNA
     intron          1066. .>2167
                     /number=11
     repeat_region   1281. .1592
                     /note="ALU repeat VI"
     repeat_region   1835. .2167
                     /note="ALU repeat VII"
BASE COUNT      477 a    626 c    485 g    579 t
ORIGIN

  Query Match             7.8%;  Score 157;  DB 9;  Length 2167;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-78;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1196 AGATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAG 1255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 AGATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAG 73

Qy       1256 CTGCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCC 1315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 CTGCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCC 133

Qy       1316 GGGTCATTGAGGACAACGAGTACACGGCTCGGGAAGG 1352
             |||||||||||||||||||||||||||||||||||||
Db        134 GGGTCATTGAGGACAACGAGTACACGGCTCGGGAAGG 170


RESULT 11
AL353092
LOCUS       AL353092               25010 bp    DNA     linear   PRI 11-FEB-2001
DEFINITION  Human DNA sequence from clone RP1-180I13 on chromosome 20 Contains
            5' end of the HCK gene for hemopoietic cell kinase (protein
            tyrosine kinase), contains ESTs, STSs, GSSs and a CpG island,
            complete sequence.
ACCESSION   AL353092
VERSION     AL353092.6  GI:9650539
KEYWORDS    HTG; CpG island; HCK; tyrosine kinase.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 25010)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 1, 2000 this sequence version replaced gi:9187765.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            IMPORTANT: This sequence is not the entire insert of clone
            RP1-180I13 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP5-836N17 is at 24911 in this sequence.
            The true right end of clone RP1-310O13 is at 100 in this sequence.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RP1-180I13 is from
            the library RPCI-1 constructed by the group of Pieter de Jong. For
            further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .25010
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP1-180I13"
                     /clone_lib="RPCI-1"
     repeat_region   1. .43
                     /note="AluJo/FLAM repeat: matches 76. .118 of consensus"
     misc_feature    complement(39. .567)
                     /note="match: GSS: Em:AQ386884"
     repeat_region   58. .175
                     /note="5S repeat: matches 1. .119 of consensus"
     repeat_region   265. .455
                     /note="MER20 repeat: matches 26. .218 of consensus"
     repeat_region   510. .586
                     /note="L2 repeat: matches 1966. .2045 of consensus"
     misc_feature    692. .906
                     /note="match: GSS: Em:AQ486790"
     repeat_region   772. .843
                     /note="2 copies 36 mer 93% conserved"
     misc_feature    1576. .2453
                     /note="CpG island"
                     /evidence=not_experimental
     gene            1846. .24191
                     /gene="HCK"
     mRNA            join(1846. .1956,21132. .21252,22789. .22831,23186. .23288,
                     24093. .>24191)
                     /gene="HCK"
                     /product="dJ180I13.1 (hemopoietic cell kinase)"
                     /note="match: cDNAs: Em:M16591 Em:X62345 Em:S74141
                     Em:J03023 Em:M16592 Em:M83666 Em:Y00487 Em:X60380
                     Em:J03579 Em:X67786 Em:AF000300 Em:AF000301 Em:AF000302
                     Em:M17031 Em:X52822 Em:X57018 Em:X15345 Em:M27454
                     Em:M57290 Em:X57684 Em:M85043 Em:M19722
                     match: ESTs: Em:AW307786"
                     /evidence=not_experimental
     misc_feature    complement(2102. .2519)
                     /note="match: GSS: Em:AQ382576"
     misc_feature    complement(2108. .2557)
                     /note="match: GSS: Em:AQ461389"
     misc_feature    complement(2196. .2570)
```

```
                     /note="match: GSS: Em:AQ240986"
repeat_region        2546. .2600
                     /note="MIR repeat: matches 89. .148 of consensus"
misc_feature         2575. .3113
                     /gene="HCK"
                     /note="match: GSS: Em:AQ201482"
misc_feature         2576. .3351
                     /gene="HCK"
                     /note="match: GSS: Em:AQ748323"
misc_feature         2594. .2974
                     /gene="HCK"
                     /note="match: GSS: Em:AQ821722"
repeat_region        2700. .2825
                     /note="MIR repeat: matches 76. .193 of consensus"
repeat_region        2826. .3131
                     /note="AluSx repeat: matches 1. .306 of consensus"
repeat_region        3132. .3189
                     /note="MIR repeat: matches 193. .252 of consensus"
repeat_region        3221. .3389
                     /note="MER5A repeat: matches 1. .184 of consensus"
repeat_region        3809. .3976
                     /note="AluSp repeat: matches 147. .313 of consensus"
repeat_region        3997. .4296
                     /note="AluY repeat: matches 3. .302 of consensus"
repeat_region        4866. .5115
                     /note="L1MC1 repeat: matches 5725. .5980 of consensus"
repeat_region        5124. .5701
                     /note="L1MB5 repeat: matches 5615. .6171 of consensus"
repeat_region        5702. .5999
                     /note="AluY repeat: matches 1. .295 of consensus"
repeat_region        6000. .6157
                     /note="L1MB5 repeat: matches 5454. .5615 of consensus"
repeat_region        6276. .6378
                     /note="L2 repeat: matches 2623. .2730 of consensus"
repeat_region        6789. .7086
                     /note="AluJb repeat: matches 1. .299 of consensus"
repeat_region        7087. .7147
                     /note="MIR repeat: matches 68. .133 of consensus"
repeat_region        7614. .7706
                     /note="MIR repeat: matches 47. .140 of consensus"
repeat_region        7720. .8014
                     /note="AluSq repeat: matches 1. .297 of consensus"
misc_feature         7895. .8073
                     /gene="HCK"
                     /note="match: STS: Em:HS336XG5"
misc_feature         8076. .8206
                     /gene="HCK"
                     /note="match: STS: Em:HS336XG5"
repeat_region        8076. .8119
                     /note="22 copies 2 mer ac 97% conserved"
repeat_region        8139. .8510
                     /note="LTR16C repeat: matches 1. .387 of consensus"
repeat_region        8750. .8941
                     /note="48 copies 4 mer cttc 62% conserved"
repeat_region        8779. .8928
                     /note="75 copies 2 mer tt 60% conserved"
repeat_region        8795. .8974
                     /note="5 copies 36 mer 68% conserved"
repeat_region        8820. .8957
                     /note="46 copies 3 mer cct 66% conserved"
repeat_region        8982. .9285
                     /note="AluJo repeat: matches 1. .304 of consensus"
repeat_region        9680. .10012
                     /note="AluSx repeat: matches 1. .311 of consensus"
repeat_region        10208. .10442
                     /note="LTR16B repeat: matches 250. .459 of consensus"
repeat_region        10443. .10751
                     /note="AluSg repeat: matches 1. .309 of consensus"
repeat_region        10752. .10991
                     /note="LTR16B repeat: matches 2. .250 of consensus"
repeat_region        11000. .11406
                     /note="L2 repeat: matches 2157. .2577 of consensus"
misc_feature         complement(11015. .11870)
                     /note="match: GSS: Em:AQ896828"
repeat_region        11407. .11707
                     /note="AluSx repeat: matches 3. .303 of consensus"
misc_feature         complement(11415. .11855)
                     /note="match: GSS: Em:AQ369535"
misc_feature         complement(11468. .11868)
                     /note="match: GSS: Em:AQ135725"
misc_feature         complement(11535. .11868)
                     /note="match: GSS: Em:AQ709950"
repeat_region        11708. .11734
                     /note="L2 repeat: matches 2132. .2157 of consensus"
misc_feature         11890. .12417
                     /gene="HCK"
                     /note="match: GSS: Em:AQ583762"
misc_feature         11920. .12309
                     /gene="HCK"
                     /note="match: GSS: Em:AQ419051"
repeat_region        12733. .13018
                     /note="AluJo repeat: matches 1. .292 of consensus"
repeat_region        13463. .13759
                     /note="AluSx repeat: matches 5. .311 of consensus"
repeat_region        13991. .14088
                     /note="MIR repeat: matches 47. .144 of consensus"
repeat_region        14544. .14681
                     /note="MIR repeat: matches 2. .144 of consensus"
repeat_region        14682. .14801
                     /note="MER5A repeat: matches 66. .187 of consensus"
repeat_region        15409. .15538
                     /note="MER5A repeat: matches 4. .160 of consensus"
repeat_region        15774. .16082
                     /note="AluSx repeat: matches 1. .310 of consensus"
repeat_region        16689. .16903
                     /note="L1MB3 repeat: matches 5953. .6176 of consensus"
repeat_region        16996. .17021
                     /note="13 copies 2 mer tt 92% conserved"

  Query Match             5.6%;  Score 112;  DB 9;  Length 25010;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-52;
  Matches  112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21142 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 21201

Qy        238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGG 289
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21202 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGG 21253


RESULT 12
MFA320181
LOCUS       MFA320181               1515 bp    mRNA    linear   PRI 06-JUL-2001
DEFINITION  Macaca fascicularis mRNA for hck protein.
ACCESSION   AJ320181
VERSION     AJ320181.1  GI:14627115
KEYWORDS    hck gene; hck protein; tyrosine kinase.
SOURCE      crab-eating macaque.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1515)
  AUTHORS   Picard,C.
  JOURNAL   Thesis (2001) Department of Experimental Oncology laboratory, U
REFERENCE   2  (bases 1 to 1515)
  AUTHORS   Picard,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2001) Picard C., U119, Inserm, bd Lei Roure,
            Marseille 13010, FRANCE
FEATURES             Location/Qualifiers
     source          1. .1515
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
```

```
     gene            1. .1515
                     /gene="hck"
     CDS             1. .1515
                     /gene="hck"
                     /function="tyrosine kinase"
                     /codon_start=1
                     /product="hck protein"
                     /protein_id="CAC44031.1"
                     /db_xref="GI:14627116"
                     /translation="MGCMKSKFLQAGGNTFSKTETSANPHCPVYVPDPTSTIKPGPNS
                     NNRNTPGIGEGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKARSLATRK
                     EGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTKGSYSL
                     SVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGSDGLCQKLSVP
                     CVSSKPQKPWEKDAWEIPRESLKLEKKLGAGQFGEVWMATYNKHTKVAVKTMKPGSMS
                     VEAFLAEANLMKTLQHDKLVKLHAVVTKEPIYIITEFMAKGSLLDFLKSDEGSKQPLP
                     KLIDFSAQIAEGMAFIEQRNYIHRDLRAANILVSASLVCKIADFGLARIIEDNEYTAR
                     EGAKFPIKWTAPEAINFGSSTIKSDVWSFGILLMEIVTYGRIPYPGMSNPEVIRALER
                     GYRMPRPENCPEELYNIMMRCWKNRPEERPTFEYIQSVLDDFYTATESQYQQQP"
BASE COUNT      388 a    424 c    419 g    284 t
ORIGIN

  Query Match             5.3%;  Score 107;  DB 9;  Length 1515;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-49;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        580 CTGGCTCCCGGCAACATGCTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGA 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CTGGCTCCCGGCAACATGCTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGA 468

Qy        640 AGCTACTCTTTGTCCGTGCGAGACTACGACCCTCGGCAGGGAGATAC 686
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 AGCTACTCTTTGTCCGTGCGAGACTACGACCCTCGGCAGGGAGATAC 515


RESULT 13
G25924
LOCUS       G25924                   366 bp    DNA     linear   STS 02-JUN-1996
DEFINITION  human STS EST50744, sequence tagged site.
ACCESSION   G25924
VERSION     G25924.1  GI:1348156
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens STSs derived from sequences in dbEST and the Unigene
            collection.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped STSs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: GATCCGAGCTCTGGAGCG
            Primer B: CCGTGTAGAAGTCATCCAGC
            STS size: 150
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                Template: 10 ng
                Primer: each 5 pM
                dNTPs: each 4 nM
                Taq Polymerase: 0.025 units/ul
                Total Vol: 20 ul

            Buffer:
                MgCl2: 1.5 mM
                KCl: 50 mM
                Tris-HCL: 10 mM
                pH: 9.3

            Derived from dbEST (genbank accession D20116).
FEATURES            Location/Qualifiers
     source         1. .366
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /map="180.1 cR from top of Chr20 linkage group"
     STS            1. .150
     primer_bind    1. .18
     primer_bind    complement(131. .150)
BASE COUNT       83 a     95 c     97 g     73 t     18 others
ORIGIN

  Query Match             4.1%;  Score 82;  DB 11;  Length 366;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-35;
  Matches   82;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1613 CCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACC 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACC 168

Qy       1673 AACAGCAGCCATGATAGGGAGG 1694
              |||||||||||||||||||||||
Db        169 AACAGCAGCCATGATAGGGAGG 190


RESULT 14
HUMHCK
LOCUS       HUMHCK                   958 bp    DNA     linear   PRI 08-NOV-1994
DEFINITION  Human hemopoietic cell kinase (HCK) gene, exon 1.
ACCESSION   M73233
VERSION     M73233.1  GI:485365
KEYWORDS    hemopoietic cell kinase.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 958)
  AUTHORS   Lichtenberg,U., Quintrell,N. and Bishop,J.M.
  TITLE     Human protein-tyrosine kinase gene HCK: expression and structural
            analysis of the promoter region
  JOURNAL   Oncogene 7 (5), 849-858 (1992)
  MEDLINE   92237010
   PUBMED   1373873
FEATURES            Location/Qualifiers
     source         1. .958
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /map="20q11-q12"
                    /cell_type="lymphocyte"
     5'UTR          1. .663
     gene           664. .958
                    /gene="HCK"
     exon           664. .807
                    /gene="HCK"
                    /product="hemopoietic cell kinase"
                    /note="G00-119-303"
                    /number=1
     intron         808. .>958
                    /gene="HCK"
                    /note="G00-119-303"
BASE COUNT      177 a    219 c    379 g    183 t
ORIGIN
```

```
  Query Match             3.8%;  Score 77;  DB 9;  Length 958;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-32;
  Matches   77;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 ACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTCAAGCTGCGAGGATCCGGGCTGCCCGCGA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 ACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTCAAGCTGCGAGGATCCGGGCTGCCCGCGA 787

Qy        149 GACGAGGAGCGGGCGCC 165
              |||||||||||||||||
Db        788 GACGAGGAGCGGGCGCC 804


RESULT 15
AB071605
LOCUS       AB071605               10348 bp    mRNA    linear   PRI 02-AUG-2002
DEFINITION  Homo sapiens mRNA for HECT domain protein LASU1, complete cds.
ACCESSION   AB071605
VERSION     AB071605.1  GI:22090625
KEYWORDS    .
SOURCE      Homo sapiens fetus brain cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Gu,J., Ren,K., Dubner,R. and Iadarola,M.J.
  TITLE     Cloning of a DNA binding protein that is a tyrosine kinase
            substrate and recognizes an upstream initiator-like sequence in the
            promoter of the preprodynorphin gene
  JOURNAL   Brain Res. Mol. Brain Res. 24 (1-4), 77-88 (1994)
  MEDLINE   95058008
   PUBMED   7968380
REFERENCE   2
  AUTHORS   Gu,J., Dubner,R., Fornace,J.A. and Iadarola,M.
  TITLE     UREB1, a tyrosine phosphorylated nuclear protein, inhibits p53
            transactivation
  JOURNAL   Oncogene 16, 2175-2178 (1995)
REFERENCE   3
  AUTHORS   Miyazaki,K., Okamoto,Y., Sakamoto,M., Kato,C., Ozaki,T.,
            Watanabe,K. and Nakagawara,A.
  TITLE     Homo sapiens LASU1 (large structure of UREB1) mRNA, complete cds
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 10348)
  AUTHORS   Nakagawara,A. and Miyazaki,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-2001) Akira Nakagawara, Chiba Cancer Center
            Reserch Institute, Division of Biochemistry; 666-2 Nitona, Chuoh-ku
            Chiba, Chiba 260-8717, Japan
            (E-mail:akiranak@chiba-cc.pref.chiba.jp, Tel:81-43-264-5431,
            Fax:81-43-265-4459)
FEATURES             Location/Qualifiers
     source          1. .10348
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp11.22"
                     /tissue_type="brain"
                     /dev_stage="fetus"
     gene            1. .10348
                     /gene="LASU1"
     5'UTR           1. .145
                     /gene="LASU1"
     CDS             146. .10228
                     /gene="LASU1"
                     /note="Homo sapiens upstream regulatory element binding
                     protein 1 (UREB1) is completely included in the 3'part of
                     this gene (9234-10160).
                     KIAA0312; HSPC272,KIAA1578,ATP synthase, H+ transporting,
                     mitochondrial F1complex, alpha subunit, isoform 1
                     large structure of UREB1 (LASU1)"
                     /codon_start=1
                     /evidence=experimental
                     /product="HECT domain protein LASU1"
                     /protein_id="BAC06833.1"
                     /db_xref="GI:22090626"
                     /translation="MVTHWLPWRQMNLLLQTLRANLKSHQQWLPELSKSSLCYQASSR
                     LGRALAELFGLLVKLCVGSPVRQRRSHHAASTTTAPTPAARSTASALTKLLTNGLSWQ
                     PPPYTPTLRFRLTFFICSVGFTSPMLFDERKYPYHLMLQKFLCSGGHNALFETFNWAL
                     SMGGKVPVSEGLEHSHLPDGTGEFLDAWLMLVEKMVNPTTVLESPHSLPAKLPGGVQN
                     FPQFSALRFLVVTQKAAFTCIKNLWNRKPLKVYGGRMAESMLAILCHILRGEPVIRER
                     LSKEKEGSRGEEDTGQEEGGSRREPQVNQQQLQQLMDMGFTREHAMEALLNTSTMEQA
                     TEYLLTHPPPIMGGVVRDLSMSEEDQMMRAIAMSLGQDIPMDQRAESPEEVACRKEEE
                     ERKAREKQEEEEAKCLEKFQDADPLEQDELHTFTDTMLPGCFHLLDELPDTVYRVCDL
                     IMTAIKRNGADYRDMILKQVVNQVWEAADVLIKAALPLTTSDTKTVSEWISQMATLPQ
                     ASNLATRILLLTLLFEELKLPCAWVVESSGILNVLIKLLEVVQPCLQAAKEQKEVQTP
                     KWITPVLLLIDFYEKTAISSKRRAQMTKYLQSNSNNWRWFDDRSGRWCSYSASNNSTI
                     DSAWKSGETSVRFTAGRRRYTVQFTTMVQVNEETGNRRPVMLTLLRVPRLNKNSKNSN
                     GQELEKTLEESKEMDIKRKENKGNDTPLALESTNTEKETSLEETKIGEILIQGLTEDM
                     VTVLIRACVSMLGVPVDPDTLHATLRLCLRLTRDHKYAMMFAELKSTRMILNLTQSSG
                     FNGFTPLVTLLLRHIIEDPCTLRHTMEKVVRSAATSGAGSTTSGVVSGSLGSREINYI
                     LRVLGPAACRNPDIFTEVANCCIRIALPAPRGSGTASDDEFENLRIKGPNAVQLVKTT
                     PLKPSPLPVIPDTIKEVIYDMLNALAAYHAPEEADKSDPKPGVMTQEVGQLLQDMGDD
                     VYQQYRSLTRQSSDFDTQSGFSINSQVFAADGASTETSASGTSQGEASTPEESRDGKK
                     DKEGDRASEEGKQKGKGSKPLMPTSTILRLLAELVRSYVGIATLIANYSYTVGQSELI
                     KEDCSVLAFVLDHLLPHTQNAEDKDTPALARLFLASLAAAGSGTDAQVALVNEVKAAL
                     GRALAMAESTEKHARLQAVMCIISTIMESCPSTSSFYSSATAKTQHNGMNNIIRLFLK
                     KGLVNDLARVPHSLDLSSPNMANTVNAALKPLETLSRIVNQPSSLFGSKSASSKNKSE
                     QDAQGASQDSSSNQQDPGEPGEAEVQEEDHDVTQTEVADGDIMDGEAETDSVVIAGQP
                     EVLSSQEMQVENELEDLIDELLERDGGSGNSTIIVSRSGEDESQEDVLMDEAPSNLSQ
                     ASTLQANREDSMNILDPEDEEEHTQEEDSSGSNEDEDDSQDEEEEEEEDEEDDQEDDE
                     GEEGDEDDDDDGSEMELDEDYPDMNASPLVRFERFDREDDLIIEFDNMFSSATDIPPS
                     PGNIPTTHPLMVRHADYSSLTLGSGSSTTRLTQGIGRSQRTLRQLTANTGHTIHVHYP
                     GNRQPNPPLILQRLLGPSAAADILQLSSSLPLQSRGRARLLVGNDDVHIIARSDDELL
                     DDFFHDQSTATSQAGTLSSIPTALTRWTEECKVLDAESMHDCVSVVKVSIVNHLEFLR
                     DEELEERREKRRKQLAEEETKITDKGKEDKENRDQSAQCTASKSNDSTEQNLSDGTPM
                     PDSYPTTPSSTDAATSESKETLGTLQSSQQQPTLPTPPALGEVPQELQSPAGEGGSST
                     QLLMPVEPEELGPTRPSGEAETTQMELSPAPTITSLSPERAEDSDALTAVSSQLEGSP
                     MDTSSLASCTLEEEAVGDTSAAGSSEQPRAGSSTPGDAPPAVAEVQGRSDGSGESAQPP
                     EDSSPPASSESSSTRDSAVAISGADSRGILEEPLPSTSSEEEDPLAGISLPEGVDPSF
                     LAALPDDIRREVLQNQLGIRPPTRTAPSTNSSAPAVVGNPGVTEVSPEFLAALPPAIQ
                     EEVLAQQRAEQQRRELAQNASSDTPMDPVTFIQTLPSDLRRSVLEDMEDSVLAVMPPD
                     IAAEAQALRREQEARQRQLMHERLFGHSSTSALSAILRSPAFTSRLSGNRGVQYTRLA
                     VQRGGTFQMGGSSSHNRPSGSNVDTLLRLRGRLLLDHEALSCLLVLLFVDEPKLNTSR
                     LHRVLRNLCYHAQTRHWVIRSLLSILQRSSESELCIETPKLTTSEEKGKKSSKSCGSS
                     SHENRPLDLLHKMESKSSNQLSWLSVSMDAALGCRTNIFQIQRSGGRKHTEKHASGGS
                     TVHIHPQAAPVVCRHVLDTLIQLAKVFPSHFTQQRTKETNCESDRERGNKACSPCSSQ
                     SSSSGICTDFWDLLVKLDNMNVSRKGKNSVKSVPVSAGGEGETSPYSLEASPLGQLMN
                     MLSHPVIRRSSLLTEKLLRLLSLISIALPENKVSEAQANSGSGASSTTTATSTTSTTT
                     TTAASTTPTPPTAPTPVTSAPALVAATAISTIVVAASTTVTTPTTATTTVSISPTTKG
                     SKSPAKVSDGGSSSTDFKMVSSGLTENQLQLSVEVLTSHSCSEEGLEDAANVLLQLSR
                     GDSGTRDTVLKLLLNGARHLGYTLCKQIGTLLAELREYNLEQQRRAQCETLSPDGLPE
                     EQPQTTKLKGKMQSRFDMAENVVIVASQKRPLGGRELQLPSMSMLTSKTSTQKFFLRV
                     LQVIIQLRDDTRRANKKAKQTGRLGSSGLGSASSIQAAVRQLEAEADAIIQMVREGQR
                     ARRQQQAATSESSQSEASVRREESPMDVDQPSPSAQDTQSIASDGTPQGEKEKEERPP
                     ELPLLSEQLSLDELWDMLGECLKELEESHDQHAVLVLQPAVEAFFLVHATERESKPPV
                     RDTRESQLAHIKDEPPPLSPAPLTPATPSSLDPFFSREPSSMHISSSLPPDTQKFLRF
                     AETHRTVLNQILRQSTTHLADGPFAVLVDYIRVLDFDVKRKYFRQELERLDEGLRKED
                     MAVHVRRDHVFEDSYRELHRKSPEEMKNRLYIVFEGEEGQDAGGLLREWYMIISREMF
                     NPMYALFRTSPGDRVTYTINPSSHCNPNHLSYFKFVGRIVAKAVYDNRLLECYFTRSF
                     YKHILGKSVRYTDMESEDYHFYQGLVYLLENDVSTLGYDLTFSTEVQEFGVCEVRDLK
                     PNGANILVTEENKKEYVHLVCQMRMTGAIRKQLAAFLEGFYEIIPKRLISIFTEQELE
                     LLISGLPTIDIDDLKSNTEYHKYQSNSIQIQWFWRALRSFDQADRAKFLQFVTGTSKV
                     PLQGFAALEGMNGIQKFQIHRDDRSTDRLPSAHTCFNQLDLPAYESFEKLRHMLLLAI
                     QECSEGFGLA"
     misc_feature    1952. .2131
                     /gene="LASU1"
                     /note="WWE domain"
     misc_feature    9305. .10225
                     /gene="LASU1"
                     /note="HECT-domain (ubiquitin-transferase)"
     3'UTR           10229. .10348
                     /gene="LASU1"
BASE COUNT     2684 a   2757 c   2656 g   2251 t
ORIGIN

  Query Match             3.8%;  Score 76;  DB 9;  Length 10348;
```

Best Local Similarity 100.0%; Pred. No. 1.8e-31;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          10 GGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGAAGACGATGA 69
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4420 GGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGAAGACGATGA 4479

Qy          70 CGACGATGGCTCTGAG 85
               ||||||||||||||||
Db        4480 CGACGATGGCTCTGAG 4495
```

Search completed: July 4, 2003, 03:58:32
Job time : 5242 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 00:36:13 ; Search time 456 Seconds
(without alignments)
9951.264 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2015	100.0	2015	24	ABK83939	Human cDNA differe
2	2015	100.0	2015	24	ABL66673	Lung cancer relate
3	1552	77.0	1926	24	ABK83940	Human cDNA differe
4	183	9.1	183	24	ABL61214	Human nucleotide f
5	181	9.0	1416	24	ABL61215	Rat/human fusion c
6	181	9.0	1542	24	ABL61216	Rat/human fusion c
7	169	8.4	369	16	AAT19957	Human gene signatu
8	133	6.6	171	22	ABA50558	Human breast cell
9	133	6.6	171	22	ABA68516	Human foetal liver
10	133	6.6	171	22	ABA35497	Probe #13963 for g
11	133	6.6	171	22	AAK16884	Human brain expres
12	133	6.6	171	22	AAK42654	Human bone marrow
13	133	6.6	171	22	AAI23408	Probe #13341 for g
14	133	6.6	171	22	AAI48728	Probe #17414 used
15	133	6.6	171	22	AAI09035	Probe #9026 used t
16	133	6.6	171	24	ABS16706	Human genome-deriv
17	113	5.6	415	22	ABA45430	Human breast cell
18	113	5.6	415	22	ABA55928	Human foetal liver
19	113	5.6	415	22	ABA25595	Probe #4061 for ge
20	113	5.6	415	22	AAK04142	Human brain expres
21	113	5.6	415	22	AAK29623	Human bone marrow
22	113	5.6	415	22	AAI14202	Probe #4135 for ge
23	113	5.6	415	22	AAI35583	Probe #4269 used t
24	113	5.6	415	22	AAI04039	Probe #4030 used t
25	113	5.6	415	24	ABS04179	Human genome-deriv
26	112	5.6	1592	20	AAZ27241	Human secreted pro
27	78	3.9	409	22	AAH99174	Human protein enco
28	77	3.8	334	21	AAA52650	Eosinophil activat
29	68	3.4	1911	24	ABK63704	Rat sequence diffe
30	66	3.3	274	22	AAK68573	Human immune/haema
c 31	65	3.2	1926	24	ABK83940	Human cDNA differe
32	31	1.5	31	22	AAI30734	Human single nucle
33	31	1.5	31	22	AAI30735	Human single nucle
34	31	1.5	31	22	AAI30736	Human single nucle
35	31	1.5	31	22	AAI30737	Human single nucle
36	31	1.5	31	22	AAI30738	Human single nucle
37	28	1.4	2298	24	ABK83935	Human cDNA differe
c 38	27	1.3	33	22	AAH41498	Human tyrosine kin
39	26	1.3	32	22	AAH41491	Human tyrosine kin
c 40	26	1.3	32	22	AAH41492	Human tyrosine kin
41	26	1.3	1602	14	AAQ46687	Chicken pp60 c-src
42	26	1.3	1759	21	AAZ29700	Wild-type chicken
43	26	1.3	1759	22	AAH28357	Nucleotide sequenc
c 44	25	1.2	32	22	AAH41501	Human tyrosine kin
45	25	1.2	51	23	ABL00375	Human silent nonco

ALIGNMENTS

RESULT 1
ABK83939
ID ABK83939 standard; cDNA; 2015 BP.
XX
AC ABK83939;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #510.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 510; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2015 BP; 512 A; 540 C; 580 G; 383 T; 0 other;

Query Match 100.0%; Score 2015; DB 24; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080

Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320

Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980

Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
                ||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
```

RESULT 2
ABL66673
ID ABL66673 standard; DNA; 2015 BP.
XX
AC ABL66673;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:5010.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 5010; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 2015 BP; 512 A; 540 C; 580 G; 383 T; 0 other;

Query Match 100.0%; Score 2015; DB 24; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGAGGCACGGAAGATGAGGAAGATGATCAGGAGGATGATGAAGGTGAAGAGGGAGATGA 60

Qy         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGATGACGACGATGGCTCTGAGGGGACCTCAGGGGCTGCCGAGCTGGGGGGGCGCTC 120

Qy        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGCGAGGATCCGGGCTGCCCGCGAGACGAGGAGCGGGCGCCAGGATGGGGTCGATG 180

Qy        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCCAGC 240

Qy        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAATAGC 300

Qy        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTTGCC 360

Qy        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAGATG 420

Qy        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAGGAG 480

Qy        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGGTTT 540

Qy        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATGCTG 600

Qy        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGCGA 660

Qy        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGACAAC 720

Qy        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCAC 780

Qy        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCCAAG 840

Qy        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTGGAG 900

Qy        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCACACC 960

Qy        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCAGAG 1020

Qy       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTCACC 1080

Qy       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTTCTG 1140

Qy       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAGATT 1200

Qy       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCTGCC 1260

Qy       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGGGTC 1320

Qy       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACAGCT 1380

Qy       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGTATC 1440

Qy       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCTGAA 1500

Qy       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAGGAG 1560

Qy       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTCGAA 1620

Qy       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAGCAG 1680

Qy       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCTCGAAGGTGGCT 1740

Qy       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCAGCACCATCCGCCAGGGCCCACACCCCCTTCCTACTCCCAGACACCCACCCTCGCTTC 1800

Qy       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCCACAGTTTCCTCATCTGTCCAGTGGGTAGGTTGGACTGGAAAATCTCTTTTTGACTC 1860

Qy       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTGCAATCCACAATCTGACATTCTCAGGAAGCCCCCAAGTTGATATTTCTATTTCCTGGA 1920

Qy       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATGGTTGGATTTTAGTTACAGCTGTGATTTGGAAGGGAAACTTTCAAAATAGTGAAATGA 1980

Qy       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
              ||||||||||||||||||||||||||||||||||
Db       1981 ATATTTAAATAAAAGATATAAATGCAAGTCTTACG 2015
```

RESULT 3
ABK83940
ID ABK83940 standard; cDNA; 1926 BP.
XX
AC ABK83940;
XX
DT 14-AUG-2002 (first entry)
XX

DE Human cDNA differentially expressed in granulocytic cells #511.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 511; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1926 BP; 497 A; 522 C; 520 G; 387 T; 0 other;

Query Match 77.0%; Score 1552; DB 24; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       178 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        85 ATGAAGTCCAAGTTCCTCCAGGTCGGAGGCAATACATTCTCAAAAACTGAAACCAGCGCC 144

Qy       238 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       145 AGCCCACACTGTCCTGTGTACGTGCCGGATCCCACATCCACCATCAAGCCGGGGCCTAAT 204

Qy       298 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       205 AGCCACAACAGCAACACACCAGGAATCAGGGAGGCAGGCTCTGAGGACATCATCGTGGTT 264

Qy       358 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       265 GCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAGCTTCCAGAAGGGGGACCAG 324

Qy       418 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       325 ATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCGATCCCTGGCCACCCGGAAG 384

Qy       478 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       385 GAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTCTCTGGAGACAGAGGAGTGG 444

Qy       538 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 597
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       445 TTTTTCAAGGGCATCAGCCGGAAGGACGCAGAGCGCCAACTGCTGGCTCCCGGCAACATG 504

Qy       598 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 657
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       505 CTGGGCTCCTTCATGATCCGGGATAGCGAGACCACTAAAGGAAGCTACTCTTTGTCCGTG 564

Qy       658 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 717
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       565 CGAGACTACGACCCTCGGCAGGGAGATACCGTGAAACATTACAAGATCCGGACCCTGGAC 624

Qy       718 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 777
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       625 AACGGGGGCTTCTACATATCCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGAC 684

Qy       778 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 837
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       685 CACTACAAGAAGGGGAACGACGGGCTCTGCCAGAAACTGTCGGTGCCCTGCATGTCTTCC 744

Qy       838 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 897
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       745 AAGCCCCAGAAGCCTTGGGAGAAAGATGCCTGGGAGATCCCTCGGGAATCCCTCAAGCTG 804

Qy       898 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 957
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       805 GAGAAGAAACTTGGAGCTGGGCAGTTTGGGGAAGTCTGGATGGCCACCTACAACAAGCAC 864

Qy       958 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 1017
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       865 ACCAAGGTGGCAGTGAAGACGATGAAGCCAGGGAGCATGTCGGTGGAGGCCTTCCTGGCA 924

Qy      1018 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 1077
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       925 GAGGCCAACGTGATGAAAACTCTGCAGCATGACAAGCTGGTCAAACTTCATGCGGTGGTC 984

Qy      1078 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1137
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       985 ACCAAGGAGCCCATCTACATCATCACGGAGTTCATGGCCAAAGGAAGCTTGCTGGACTTT 1044

Qy      1138 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1197
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1045 CTGAAAAGTGATGAGGGCAGCAAGCAGCCATTGCCAAAACTCATTGACTTCTCAGCCCAG 1104

Qy      1198 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1257
```

```
Db       1105 ATTGCAGAAGGCATGGCCTTCATCGAGCAGAGGAACTACATCCACCGAGACCTCCGAGCT 1164

Qy       1258 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 GCCAACATCTTGGTCTCTGCATCCCTGGTGTGTAAGATTGCTGACTTTGGCCTGGCCCGG 1224

Qy       1318 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1225 GTCATTGAGGACAACGAGTACACGGCTCGGGAAGGGGCCAAGTTCCCCATCAAGTGGACA 1284

Qy       1378 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 GCTCCTGAAGCCATCAACTTTGGCTCCTTCACCATCAAGTCAGACGTCTGGTCCTTTGGT 1344

Qy       1438 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1345 ATCCTGCTGATGGAGATCGTCACCTACGGCCGGATCCCTTACCCAGGGATGTCAAACCCT 1404

Qy       1498 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1405 GAAGTGATCCGAGCTCTGGAGCGTGGATACCGGATGCCTCGCCCAGAGAACTGCCCAGAG 1464

Qy       1558 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1465 GAGCTCTACAACATCATGATGCGCTGCTGGAAAAACCGTCCGGAGGAGCGGCCGACCTTC 1524

Qy       1618 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1525 GAATACATCCAGAGTGTGCTGGATGACTTCTACACGGCCACAGAGAGCCAGTACCAACAG 1584

Qy       1678 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1585 CAGCCATGATAGGGAGGACCAGGGCAGGGCAGGGGGTGCCCAGGTGGTGGCT 1636
```

RESULT 4
ABL61214
ID ABL61214 standard; DNA; 183 BP.
XX
AC ABL61214;
XX
DT 04-SEP-2002 (first entry)
XX
DE Human nucleotide fragment capable of inactivating HIV Nef protein.
XX
KW Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW pathogenicity; diagnosis; AIDS; human; ds.
XX
OS Homo sapiens.
XX
PN DE10109532-C1.
XX
PD 13-JUN-2002.
XX
PF 28-FEB-2001; 2001DE-1009532.
XX
PR 28-FEB-2001; 2001DE-1009532.
XX
PA (GEYE/) GEYER M.
PA (FACK/) FACKLER O.
XX
PI Geyer M;
XX
DR WPI; 2002-418264/45.
XX
PT New fusion protein that blocks Nef protein, useful for treatment or
PT diagnosis of acquired immune deficiency syndrome, has high specificity
PT and affinity -
XX
PS Claim 12; Page 14; 22pp; German.
XX
CC This invention describes a novel fusion protein for blocking the Nef
CC protein of human immune deficiency virus (HIV) which comprises: (i)
CC protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC linker between protein domains 1 and 2. The products of the invention
CC have virucide and anti-HIV activity and are capable of neutralising Nef,
CC an accessory protein essential for pathogenicity of HIV-1. The fusion
CC protein of the invention comprises the LL domain of the beta-subunit of
CC the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC of the invention are used for in vitro diagnosis of AIDS and for
CC treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC unlike HIV protease, has no human homologue, so the fusion protein (which
CC binds Nef with very high affinity) should cause essentially no side
CC effects. This sequence represents a human derived nucleotide fragment
CC used in the construction of the fusion protein of the invention and which
CC contains a PXXP-motif binding motif useful to the invention.
XX
SQ Sequence 183 BP; 41 A; 50 C; 56 G; 36 T; 0 other;

```
  Query Match             9.1%;  Score 183;  DB 24;  Length 183;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-79;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        337 TCTGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTC 60

Qy        397 AGCTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCT 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCT 120

Qy        457 CGATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGAC 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGAC 180

Qy        517 TCT 519
              |||
Db        181 TCT 183
```

RESULT 5
ABL61215
ID ABL61215 standard; DNA; 1416 BP.
XX
AC ABL61215;
XX
DT 04-SEP-2002 (first entry)
XX
DE Rat/human fusion construct capable of inactivating HIV Nef protein.
XX
KW Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW pathogenicity; diagnosis; AIDS; rat; human; ds.
XX
OS Rattus sp.
OS Homo sapiens.
OS Synthetic.
XX
PN DE10109532-C1.
XX
PD 13-JUN-2002.
XX
PF 28-FEB-2001; 2001DE-1009532.
XX
PR 28-FEB-2001; 2001DE-1009532.
XX
PA (GEYE/) GEYER M.
PA (FACK/) FACKLER O.
XX
PI Geyer M;
XX
DR WPI; 2002-418264/45.
XX
PT New fusion protein that blocks Nef protein, useful for treatment or

```
PT   diagnosis of acquired immune deficiency syndrome, has high specificity
PT   and affinity -
XX
PS   Claim 13; Page 14-15; 22pp; German.
XX
CC   This invention describes a novel fusion protein for blocking the Nef
CC   protein of human immune deficiency virus (HIV) which comprises: (i)
CC   protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC   protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC   linker between protein domains 1 and 2. The products of the invention
CC   have virucide and anti-HIV activity and are capable of neutralising Nef,
CC   an accessory protein essential for pathogenicity of HIV-1. The fusion
CC   protein of the invention comprises the LL domain of the beta-subunit of
CC   the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC   tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC   of the invention are used for in vitro diagnosis of AIDS and for
CC   treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC   unlike HIV protease, has no human homologue, so the fusion protein (which
CC   binds Nef with very high affinity) should cause essentially no side
CC   effects. This sequence represents a fusion construct composed of a rat
CC   nucleotide fragment which contains a dileucine (LL) motif and a human
CC   nucleotide fragment containing a PXXP-motif binding domain useful to the
CC   invention.
XX
SQ   Sequence 1416 BP; 340 A; 383 C; 386 G; 307 T; 0 other;

  Query Match             9.0%;  Score 181;  DB 24;  Length 1416;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-78;
  Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1233 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 1292

Qy        399 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1293 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 1352

Qy        459 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 1412

Qy        519 T 519
              |
Db       1413 T 1413


RESULT 6
ABL61216
ID   ABL61216 standard; DNA; 1542 BP.
XX
AC   ABL61216;
XX
DT   04-SEP-2002  (first entry)
XX
DE   Rat/human fusion construct capable of inactivating HIV Nef protein.
XX
KW   Nef protein; fusion protein; virucide; anti-HIV; accessory protein;
KW   pathogenicity; diagnosis; AIDS; rat; human; ds.
XX
OS   Rattus sp.
OS   Homo sapiens.
OS   Synthetic.
XX
PN   DE10109532-C1.
XX
PD   13-JUN-2002.
XX
PF   28-FEB-2001; 2001DE-1009532.
XX
PR   28-FEB-2001; 2001DE-1009532.
XX
PA   (GEYE/) GEYER M.
PA   (FACK/) FACKLER O.
XX
PI   Geyer M;
XX
DR   WPI; 2002-418264/45.
XX
PT   New fusion protein that blocks Nef protein, useful for treatment or
PT   diagnosis of acquired immune deficiency syndrome, has high specificity
PT   and affinity -
XX
PS   Claim 16; Page 15-16; 22pp; German.
XX
CC   This invention describes a novel fusion protein for blocking the Nef
CC   protein of human immune deficiency virus (HIV) which comprises: (i)
CC   protein domain 1 that binds to a di-leucine (LL) motif; (ii) a
CC   protein domain 2 that binds to a PxxP motif; and (iii) a polypeptide
CC   linker between protein domains 1 and 2. The products of the invention
CC   have virucide and anti-HIV activity and are capable of neutralising Nef,
CC   an accessory protein essential for pathogenicity of HIV-1. The fusion
CC   protein of the invention comprises the LL domain of the beta-subunit of
CC   the adapter-protein complex AP-1 and the PxxP binding SH3 domain of
CC   tyrosine kinase Hck, linked through a 60 amino acid peptide. The products
CC   of the invention are used for in vitro diagnosis of AIDS and for
CC   treatment of AIDS. The LL and PxxP motifs are specific for Nef, which,
CC   unlike HIV protease, has no human homologue, so the fusion protein (which
CC   binds Nef with very high affinity) should cause essentially no side
CC   effects. This sequence represents a fusion construct composed of a rat
CC   nucleotide fragment which contains a dileucine (LL) motif and a human
CC   nucleotide fragment containing a PXXP-motif binding domain useful to the
CC   invention.
XX
SQ   Sequence 1542 BP; 369 A; 419 C; 427 G; 327 T; 0 other;

  Query Match             9.0%;  Score 181;  DB 24;  Length 1542;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-78;
  Matches  181;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1290 TGAGGACATCATCGTGGTTGCCCTGTATGATTACGAGGCCATTCACCACGAAGACCTCAG 1349

Qy        399 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 CTTCCAGAAGGGGGACCAGATGGTGGTCCTAGAGGAATCCGGGGAGTGGTGGAAGGCTCG 1409

Qy        459 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1410 ATCCCTGGCCACCCGGAAGGAGGGCTACATCCCAAGCAACTATGTCGCCCGCGTTGACTC 1469

Qy        519 T 519
              |
Db       1470 T 1470


RESULT 7
AAT19957
ID   AAT19957 standard; cDNA to mRNA; 369 BP.
XX
AC   AAT19957;
XX
DT   17-JUL-1996  (first entry)
XX
DE   Human gene signature HUMGS01089.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
```

```
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 520; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 369 BP; 82 A; 97 C; 102 G; 75 T; 13 other;

  Query Match             8.4%;  Score 169;  DB 16;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-72;
  Matches  169;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1535 CTCGCCCAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACC 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 CTCGCCCAGAGAACTGCCCAGAGGAGCTCTACAACATCATGATGCGCTGCTGGAAAAACC 92

Qy       1595 GTCCGGAGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGG 1654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GTCCGGAGGAGCGGCCGACCTTCGAATACATCCAGAGTGTGCTGGATGACTTCTACACGG 152

Qy       1655 CCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCA 1703
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 CCACAGAGAGCCAGTACCAACAGCAGCCATGATAGGGAGGACCAGGGCA 201


RESULT 8
ABA50558
ID   ABA50558 standard; DNA; 171 BP.
XX
AC   ABA50558;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9253.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes -
XX
PS   Claim 4; SEQ ID NO 9253; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 9
ABA68516
ID   ABA68516 standard; DNA; 171 BP.
XX
AC   ABA68516;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16821.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
```

```
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 16821; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 10
ABA35497
ID   ABA35497 standard; DNA; 171 BP.
XX
AC   ABA35497;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #13963 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 13963; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 11
AAK16884
ID   AAK16884 standard; DNA; 171 BP.
XX
AC   AAK16884;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16875.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
```

```
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 16875; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 12
AAK42654
ID   AAK42654 standard; DNA; 171 BP.
XX
AC   AAK42654;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 17211.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 17211; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 13
AAI23408
ID   AAI23408 standard; DNA; 171 BP.
XX
AC   AAI23408;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #13341 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
```

```
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 13341; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 14
AAI48728
ID   AAI48728 standard; DNA; 171 BP.
XX
AC   AAI48728;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #17414 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 17414; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

  Query Match             6.6%;  Score 133;  DB 22;  Length 171;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133


RESULT 15
AAI09035
ID   AAI09035 standard; DNA; 171 BP.
XX
AC   AAI09035;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #9026 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 9026; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
```

CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 171 BP; 35 A; 53 C; 46 G; 37 T; 0 other;

Query Match 6.6%; Score 133; DB 22; Length 171;
Best Local Similarity 100.0%; Pred. No. 7.1e-55;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1352 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCAAGTTCCCCATCAAGTGGACAGCTCCTGAAGCCATCAACTTTGGCTCCTTCACCA 60

Qy       1412 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTCAGACGTCTGGTCCTTTGGTATCCTGCTGATGGAGATCGTCACCTACGGCCGGA 120

Qy       1472 TCCCTTACCCAGG 1484
             |||||||||||||
Db        121 TCCCTTACCCAGG 133
```

Search completed: July 4, 2003, 02:31:01
Job time : 458 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 04:47:18 ; Search time 458 Seconds
(without alignments)
9907.809 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31	1.5	31	22	AAI30734	Human single nucle
	2	31	1.5	31	22	AAI30735	Human single nucle
	3	31	1.5	31	22	AAI30736	Human single nucle
	4	31	1.5	31	22	AAI30737	Human single nucle
	5	31	1.5	31	22	AAI30738	Human single nucle
c	6	27	1.3	33	22	AAH41498	Human tyrosine kin
	7	26	1.3	32	22	AAH41491	Human tyrosine kin
c	8	26	1.3	32	22	AAH41492	Human tyrosine kin
c	9	25	1.2	32	22	AAH41501	Human tyrosine kin
	10	25	1.2	51	23	ABL00375	Human silent nonco
	11	24	1.2	32	22	AAH41500	Human tyrosine kin
	12	24	1.2	78	22	AAC90044	PCR primer used to
	13	21	1.0	21	22	AAF95624	Human gene single
	14	21	1.0	21	22	AAF95625	Human gene single
	15	21	1.0	21	22	AAF95626	Human gene single
	16	21	1.0	21	22	AAF95627	Human gene single
	17	21	1.0	21	22	AAF95628	Human gene single
	18	21	1.0	21	22	AAF95629	Human gene single
	19	21	1.0	21	22	AAF95630	Human gene single
	20	20	1.0	20	16	AAT41207	Human gene signatu
c	21	20	1.0	20	16	AAT41208	Human gene signatu
c	22	19	0.9	51	22	AAL33024	Human SNP oligonuc
c	23	19	0.9	51	22	AAL33025	Human SNP oligonuc
	24	18	0.9	19	21	AAA82879	cdk4 ribozyme bind
	25	18	0.9	19	22	AAH58041	Cell-cycle depende
	26	18	0.9	20	20	AAX29342	Chemically modifie
c	27	18	0.9	20	20	AAX29331	JNK2-specific prob
c	28	18	0.9	20	21	AAC62874	JNK antisense olig
	29	18	0.9	20	21	AAC62885	JNK antisense olig
c	30	18	0.9	20	21	AAA48651	Antisense oligonuc
c	31	18	0.9	20	22	AAH23754	JNK1 antisense oli
c	32	18	0.9	20	22	AAF99183	Immunostimulatory
c	33	18	0.9	20	24	ABL39057	Immunostimulatory
	34	18	0.9	34	22	AAH41497	Human tyrosine kin
c	35	18	0.9	48	24	ABK30196	CYP2D6 gene polymo
c	36	18	0.9	51	24	ABK30195	CYP2D6 gene polymo
	37	17	0.8	19	21	AAA82878	cdk4 ribozyme bind
	38	17	0.8	19	22	AAH58040	Cell-cycle depende
c	39	17	0.8	20	20	AAX01356	PCR primer for mou
c	40	17	0.8	20	22	AAF72970	Human daxx inhibit
c	41	17	0.8	21	24	ABK40441	Forward PCR primer
c	42	17	0.8	23	14	AAQ49744	pTK primer pTK2.
c	43	17	0.8	23	16	AAT03086	Protein tyrosine-k
	44	17	0.8	25	21	AAZ37264	PCR primer for SGR
c	45	17	0.8	57	22	AAS04769	Synthetic gene shG

ALIGNMENTS

RESULT 1
AAI30734
ID AAI30734 standard; DNA; 31 BP.
XX
AC AAI30734;
XX
DT 18-OCT-2001 (first entry)
XX
DE Human single nucleotide polymorphism (SNP) HCK 1.
XX
KW Human; resequence; genotype; disease; forensic; paternity testing;
KW single nucleotide polymorphism; SNP; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(16,T)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200166800-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001WO-US07268.
XX
PR 07-MAR-2000; 2000US-0187510.
PR 22-MAY-2000; 2000US-0206129.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Cargill M, Ireland JS, Lander ES;

```
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 104; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 10 A; 9 C; 7 G; 5 T; 0 other;

  Query Match             1.5%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1152 GGGCAGCAAGCAGCCATTGCCAAAACTCATT 1182
             |||||||||||||||||||||||||||||||
Db          1 GGGCAGCAAGCAGCCATTGCCAAAACTCATT 31


RESULT 2
AAI30735
ID   AAI30735 standard; DNA; 31 BP.
XX
AC   AAI30735;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) HCK 2.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 104; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 8 A; 8 C; 9 G; 6 T; 0 other;

  Query Match             1.5%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1195 CAGATTGCAGAAGGCATGGCCTTCATCGAGC 1225
             |||||||||||||||||||||||||||||||
Db          1 CAGATTGCAGAAGGCATGGCCTTCATCGAGC 31


RESULT 3
AAI30736
ID   AAI30736 standard; DNA; 31 BP.
XX
AC   AAI30736;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) HCK 3.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,A)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 104; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
```

```
SQ   Sequence 31 BP; 6 A; 9 C; 6 G; 10 T; 0 other;

  Query Match             1.5%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1405 TTCACCATCAAGTCAGACGTCTGGTCCTTTG 1435
              |||||||||||||||||||||||||||||||
Db          1 TTCACCATCAAGTCAGACGTCTGGTCCTTTG 31


RESULT 4
AAI30737
ID   AAI30737 standard; DNA; 31 BP.
XX
AC   AAI30737;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) HCK 4.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,A)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 104; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 7 A; 7 C; 10 G; 7 T; 0 other;

  Query Match             1.5%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        921 GTTTGGGGAAGTCTGGATGGCCACCTACAAC 951
              |||||||||||||||||||||||||||||||
Db          1 GTTTGGGGAAGTCTGGATGGCCACCTACAAC 31


RESULT 5
AAI30738
ID   AAI30738 standard; DNA; 31 BP.
XX
AC   AAI30738;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) HCK 5.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 104; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 12 A; 11 C; 5 G; 3 T; 0 other;

  Query Match             1.5%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        219 AAAAACTGAAACCAGCGCCAGCCCACACTGT 249
              |||||||||||||||||||||||||||||||
Db          1 AAAAACTGAAACCAGCGCCAGCCCACACTGT 31


RESULT 6
AAH41498/c
ID   AAH41498 standard; DNA; 33 BP.
XX
AC   AAH41498;
XX
DT   23-AUG-2001  (first entry)
XX
DE   Human tyrosine kinase Hck PCR primer SEQ ID NO:10.
```

XX
KW Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW Hck signal transduction; human immunodeficiency virus; HIV infection;
KW anticancer; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200132869-A1.
XX
PD 10-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-JP07500.
XX
PR 29-OCT-1999; 99JP-0309957.
XX
PA (SSSE) SSP CO LTD.
XX
PI Taniyama T, Narita T;
XX
DR WPI; 2001-316440/33.
XX
PT New proteins which bind to human tyrosine kinase Hck for promotion of
PT apoptosis and for the elucidation of the mechanism of Hck signal
PT transduction -
XX
PS Example 3; Page 33; 45pp; Japanese.
XX
CC The present invention describes a protein, designated HSB-1, which binds
CC to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC encoding the protein and its derivatives; (2) recombinant vectors
CC containing the nucleic acids; and (3) host cells transformed by the
CC vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC of Hck signal transduction and of the role of Hck in human
CC immunodeficiency virus (HIV) infection. They can be used for the
CC treatment of infections and other diseases with which Hck is associated.
CC They promote the anticancer activity of tumour necrosis factor alpha.
CC The present sequence represents a PCR primer for the human tyrosine
CC kinase Hck, which is used in an example from the present invention.
XX
SQ Sequence 33 BP; 2 A; 8 C; 11 G; 12 T; 0 other;

Query Match 1.3%; Score 27; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1657 ACAGAGAGCCAGTACCAACAGCAGCCA 1683
||||||||||||||||||||||||||||
Db 33 ACAGAGAGCCAGTACCAACAGCAGCCA 7

RESULT 7
AAH41491
ID AAH41491 standard; DNA; 32 BP.
XX
AC AAH41491;
XX
DT 23-AUG-2001 (first entry)
XX
DE Human tyrosine kinase Hck binding protein cloning PCR primer SEQ:3.
XX
KW Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW Hck signal transduction; human immunodeficiency virus; HIV infection;
KW anticancer; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200132869-A1.
XX
PD 10-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-JP07500.
XX
PR 29-OCT-1999; 99JP-0309957.
XX
PA (SSSE) SSP CO LTD.
XX
PI Taniyama T, Narita T;
XX
DR WPI; 2001-316440/33.
XX
PT New proteins which bind to human tyrosine kinase Hck for promotion of
PT apoptosis and for the elucidation of the mechanism of Hck signal
PT transduction -
XX
PS Example 1; Page 30; 45pp; Japanese.
XX
CC The present invention describes a protein, designated HSB-1, which binds
CC to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC encoding the protein and its derivatives; (2) recombinant vectors
CC containing the nucleic acids; and (3) host cells transformed by the
CC vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC of Hck signal transduction and of the role of Hck in human
CC immunodeficiency virus (HIV) infection. They can be used for the
CC treatment of infections and other diseases with which Hck is associated.
CC They promote the anticancer activity of tumour necrosis factor alpha.
CC The present sequence represents a PCR primer used in the cloning of
CC HSB-1, which is used in an example from the present invention.
XX
SQ Sequence 32 BP; 8 A; 5 C; 9 G; 10 T; 0 other;

Query Match 1.3%; Score 26; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 350 TCGTGGTTGCCCTGTATGATTACGAG 375
|||||||||||||||||||||||||
Db 7 TCGTGGTTGCCCTGTATGATTACGAG 32

RESULT 8
AAH41492/c
ID AAH41492 standard; DNA; 32 BP.
XX
AC AAH41492;
XX
DT 23-AUG-2001 (first entry)
XX
DE Human tyrosine kinase Hck binding protein cloning PCR primer SEQ:4.
XX
KW Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW Hck signal transduction; human immunodeficiency virus; HIV infection;
KW anticancer; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200132869-A1.
XX
PD 10-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-JP07500.
XX
PR 29-OCT-1999; 99JP-0309957.
XX
PA (SSSE) SSP CO LTD.
XX
PI Taniyama T, Narita T;
XX
DR WPI; 2001-316440/33.
XX

PT New proteins which bind to human tyrosine kinase Hck for promotion of
PT apoptosis and for the elucidation of the mechanism of Hck signal
PT transduction -
XX
PS Example 1; Page 31; 45pp; Japanese.
XX
CC The present invention describes a protein, designated HSB-1, which binds
CC to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC encoding the protein and its derivatives; (2) recombinant vectors
CC containing the nucleic acids; and (3) host cells transformed by the
CC vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC of Hck signal transduction and of the role of Hck in human
CC immunodeficiency virus (HIV) infection. They can be used for the
CC treatment of infections and other diseases with which Hck is associated.
CC They promote the anticancer activity of tumour necrosis factor alpha.
CC The present sequence represents a PCR primer used in the cloning of
CC HSB-1, which is used in an example from the present invention.
XX
SQ Sequence 32 BP; 7 A; 10 C; 9 G; 6 T; 0 other;

Query Match 1.3%; Score 26; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 796 GACGGGCTCTGCCAGAAACTGTCGGT 821
||||||||||||||||||||||||||
Db 32 GACGGGCTCTGCCAGAAACTGTCGGT 7

RESULT 9
AAH41501/c
ID AAH41501 standard; DNA; 32 BP.
XX
AC AAH41501;
XX
DT 23-AUG-2001 (first entry)
XX
DE Human tyrosine kinase Hck binding protein cloning PCR primer SEQ:15.
XX
KW Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW Hck signal transduction; human immunodeficiency virus; HIV infection;
KW anticancer; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200132869-A1.
XX
PD 10-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-JP07500.
XX
PR 29-OCT-1999; 99JP-0309957.
XX
PA (SSSE) SSP CO LTD.
XX
PI Taniyama T, Narita T;
XX
DR WPI; 2001-316440/33.
XX
PT New proteins which bind to human tyrosine kinase Hck for promotion of
PT apoptosis and for the elucidation of the mechanism of Hck signal
PT transduction -
XX
PS Example 4; Page 41; 45pp; Japanese.
XX
CC The present invention describes a protein, designated HSB-1, which binds
CC to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC encoding the protein and its derivatives; (2) recombinant vectors
CC containing the nucleic acids; and (3) host cells transformed by the
CC vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC of Hck signal transduction and of the role of Hck in human
CC immunodeficiency virus (HIV) infection. They can be used for the
CC treatment of infections and other diseases with which Hck is associated.
CC They promote the anticancer activity of tumour necrosis factor alpha.
CC The present sequence represents a PCR primer used in the cloning of
CC HSB-1, which is used in an example from the present invention.
XX
SQ Sequence 32 BP; 8 A; 9 C; 10 G; 5 T; 0 other;

Query Match 1.2%; Score 25; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 505 GCCCGCGTTGACTCTCTGGAGACAG 529
|||||||||||||||||||||||||
Db 32 GCCCGCGTTGACTCTCTGGAGACAG 8

RESULT 10
ABL00375
ID ABL00375 standard; DNA; 51 BP.
XX
AC ABL00375;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:366.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
PR 24-NOV-1999; 99US-0167383.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
PS Claim 1; Page 359; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the

```
CC   polymorphic protein within appropriate physiological samples).
XX
SQ   Sequence 51 BP; 10 A; 17 C; 13 G; 11 T; 0 other;

  Query Match             1.2%;  Score 25;  DB 23;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 0.09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        722 GGGGCTTCTACATATCCCCCCGAAG 746
              |||||||||||||||||||||||||
Db          1 GGGGCTTCTACATATCCCCCCGAAG 25


RESULT 11
AAH41500
ID   AAH41500 standard; DNA; 32 BP.
XX
AC   AAH41500;
XX
DT   23-AUG-2001  (first entry)
XX
DE   Human tyrosine kinase Hck binding protein cloning PCR primer SEQ:14.
XX
KW   Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW   tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW   Hck signal transduction; human immunodeficiency virus; HIV infection;
KW   anticancer; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200132869-A1.
XX
PD   10-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-JP07500.
XX
PR   29-OCT-1999;   99JP-0309957.
XX
PA   (SSSE ) SSP CO LTD.
XX
PI   Taniyama T,  Narita T;
XX
DR   WPI; 2001-316440/33.
XX
PT   New proteins which bind to human tyrosine kinase Hck for promotion of
PT   apoptosis and for the elucidation of the mechanism of Hck signal
PT   transduction  -
XX
PS   Example 4; Page 41; 45pp; Japanese.
XX
CC   The present invention describes a protein, designated HSB-1, which binds
CC   to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC   encoding the protein and its derivatives; (2) recombinant vectors
CC   containing the nucleic acids; and (3) host cells transformed by the
CC   vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC   tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC   apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC   of Hck signal transduction and of the role of Hck in human
CC   immunodeficiency virus (HIV) infection. They can be used for the
CC   treatment of infections and other diseases with which Hck is associated.
CC   They promote the anticancer activity of tumour necrosis factor alpha.
CC   The present sequence represents a PCR primer used in the cloning of
CC   HSB-1, which is used in an example from the present invention.
XX
SQ   Sequence 32 BP; 9 A; 4 C; 10 G; 9 T; 0 other;

  Query Match             1.2%;  Score 24;  DB 22;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 0.28;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        529 GAGGAGTGGTTTTTCAAGGGCATC 552
              ||||||||||||||||||||||||
Db          9 GAGGAGTGGTTTTTCAAGGGCATC 32


RESULT 12
AAC90044
ID   AAC90044 standard; DNA; 78 BP.
XX
AC   AAC90044;
XX
DT   13-MAR-2001  (first entry)
XX
DE   PCR primer used to create a library of RRT-Hck SH3 domains.
XX
KW   SH3 domain; human; Src homology region 3 domain; RT-loop; Hck protein;
KW   PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200072742-A2.
XX
PD   07-DEC-2000.
XX
PF   26-MAY-2000; 2000WO-FI00477.
XX
PR   26-MAY-1999;   99US-0136085.
XX
PA   (SAKS/) SAKSELA K.
XX
PI   Saksela K,  Hiipakka M;
XX
DR   WPI; 2001-061424/07.
XX
PT   A method for generating Src homology region 3 (SH3) domains with
PT   tailored binding properties or artificial SH3 domains, comprises
PT   employing random manipulation of the SH3 RT-loop sequence -
XX
PS   Example 1; Page 10; 34pp; English.
XX
CC   The present invention relates to a method for generating Src homology
CC   region 3 (SH3) domains with tailored binding properties. The method
CC   comprises producing a collection of SH3 domains containing a randomised
CC   RT-loop (RRT-SH3 domains). Human p59 Hck was used in the present
CC   invention as the SH3 domain. The present sequence is a PCR primer, which
CC   was used to create a library of RRT-Hck SH3 domains. The generated SH3
CC   domains are useful for inhibiting, activating or modifying the functions
CC   of cellular or pathogen-encoded proteins for research or therapeutic
CC   purposes.
XX
SQ   Sequence 78 BP; 13 A; 13 C; 17 G; 17 T; 18 other;

  Query Match             1.2%;  Score 24;  DB 22;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 0.28;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        391 GACCTCAGCTTCCAGAAGGGGGAC 414
              ||||||||||||||||||||||||
Db         55 GACCTCAGCTTCCAGAAGGGGGAC 78


RESULT 13
AAF95624
ID   AAF95624 standard; DNA; 21 BP.
XX
AC   AAF95624;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #385.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 75; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 3 A; 7 C; 6 G; 5 T; 0 other;

  Query Match             1.0%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 8.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        507 CCGCGTTGACTCTCTGGAGAC 527
              |||||||||||||||||||||
Db          1 CCGCGTTGACTCTCTGGAGAC 21


RESULT 14
AAF95625
ID   AAF95625 standard; DNA; 21 BP.
XX
AC   AAF95625;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #386.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 75; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 8 A; 4 C; 7 G; 2 T; 0 other;

  Query Match             1.0%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 8.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        773 TGGACCACTACAAGAAGGGGA 793
              |||||||||||||||||||||
Db          1 TGGACCACTACAAGAAGGGGA 21


RESULT 15
AAF95626
ID   AAF95626 standard; DNA; 21 BP.
XX
AC   AAF95626;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #387.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
```

XX
PD 15-MAR-2001.
XX
PF 07-SEP-2000; 2000WO-US24503.
XX
PR 10-SEP-1999; 99US-0153357.
PR 26-JUL-2000; 2000US-0220947.
PR 16-AUG-2000; 2000US-0225724.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;
XX
DR WPI; 2001-226749/23.
XX
PT Nucleic acids comprising single nucleotide polymorphisms, useful in
PT applications such as forensics, paternity testing, medicine, genetic
PT analysis and phenotype correlations to diseases such as diabetes and
PT atherosclerosis -
XX
PS Examples; Page 75; 242pp; English.
XX
CC The present invention provides a method of diagnosing a vascular disease
CC in an individual, involving determining the sequence at various
CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC genes. The sequences at a number of polymorphic sites are also provided
CC in the specification. In particular, the method can be used in the
CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC disease, stroke, peripheral vascular diseases, venous thromboembolism
CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC useful in forensics, paternity testing, genetic analysis and phenotype
CC correlations to diseases. The present sequence is an example of one of
CC the human gene SNPS shown in the specification.
XX
SQ Sequence 21 BP; 2 A; 5 C; 6 G; 8 T; 0 other;

Query Match 1.0%; Score 21; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      347 TCATCGTGGTTGCCCTGTATG 367
            |||||||||||||||||||||
Db        1 TCATCGTGGTTGCCCTGTATG 21
```

Search completed: July 4, 2003, 07:04:56
Job time : 459 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 06:57:16 ; Search time 98 Seconds
(without alignments)
6305.650 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18	0.9	20	2	US-08-910-629A-31	Sequence 31, Appl
	2	18	0.9	20	2	US-08-910-629A-42	Sequence 42, Appl
c	3	18	0.9	20	3	US-09-209-668-7	Sequence 7, Appli
c	4	18	0.9	20	3	US-09-287-796-31	Sequence 31, Appl
	5	18	0.9	20	3	US-09-287-796-42	Sequence 42, Appl
c	6	18	0.9	20	4	US-09-130-616-31	Sequence 31, Appl
	7	18	0.9	20	4	US-09-130-616-42	Sequence 42, Appl
c	8	17	0.8	20	2	US-08-730-876-2	Sequence 2, Appli
c	9	17	0.8	20	4	US-09-490-692-71	Sequence 71, Appl
c	10	17	0.8	23	1	US-08-222-616-2	Sequence 2, Appli
c	11	17	0.8	23	4	US-08-446-648-2	Sequence 2, Appli
c	12	17	0.8	23	5	PCT-US95-04228-2	Sequence 2, Appli
c	13	16	0.8	20	4	US-09-506-073-82	Sequence 82, Appl
	14	16	0.8	24	2	US-08-859-998-598	Sequence 598, App
	15	16	0.8	24	4	US-09-225-928-598	Sequence 598, App
c	16	15	0.7	18	3	US-08-951-923-51	Sequence 51, Appl
c	17	15	0.7	18	4	US-08-584-040-6218	Sequence 6218, Ap
	18	15	0.7	19	1	US-08-400-580A-11	Sequence 11, Appl
	19	15	0.7	31	2	US-08-942-423-51	Sequence 51, Appl
	20	15	0.7	36	3	US-08-951-923-52	Sequence 52, Appl
c	21	15	0.7	36	3	US-08-724-586-3	Sequence 3, Appli
c	22	15	0.7	36	4	US-09-421-632-3	Sequence 3, Appli
c	23	15	0.7	36	4	US-09-932-190-3	Sequence 3, Appli
c	24	15	0.7	45	2	US-08-039-198B-3	Sequence 3, Appli
	25	15	0.7	72	2	US-08-707-237A-47	Sequence 47, Appl
	26	14	0.7	17	4	US-08-584-040-7661	Sequence 7661, Ap
c	27	14	0.7	18	1	US-08-105-483-197	Sequence 197, App
c	28	14	0.7	18	1	US-08-220-151-78	Sequence 78, Appl
c	29	14	0.7	18	1	US-08-413-118-78	Sequence 78, Appl
c	30	14	0.7	18	1	US-08-224-657-54	Sequence 54, Appl
c	31	14	0.7	18	1	US-08-709-209-197	Sequence 197, App
c	32	14	0.7	18	1	US-08-458-101-197	Sequence 197, App
c	33	14	0.7	18	2	US-08-184-009-52	Sequence 52, Appl
	34	14	0.7	18	2	US-08-173-489C-11	Sequence 11, Appl
c	35	14	0.7	18	2	US-08-417-210A-52	Sequence 52, Appl
c	36	14	0.7	18	2	US-08-585-684B-2737	Sequence 2737, Ap
c	37	14	0.7	18	2	US-08-458-356-52	Sequence 52, Appl
c	38	14	0.7	18	3	US-08-473-446-78	Sequence 78, Appl
c	39	14	0.7	18	4	US-09-038-073-2737	Sequence 2737, Ap
c	40	14	0.7	18	4	US-08-460-736-52	Sequence 52, Appl
c	41	14	0.7	18	4	US-09-354-138-54	Sequence 54, Appl
	42	14	0.7	20	1	US-08-639-763-6	Sequence 6, Appli
c	43	14	0.7	20	4	US-09-270-542-161	Sequence 161, App
c	44	14	0.7	20	4	US-09-798-096-87	Sequence 87, Appl
c	45	14	0.7	21	1	US-08-056-200-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-08-910-629A-31/c
; Sequence 31, Application US/08910629A
; Patent No. 5877309
; GENERAL INFORMATION:
; APPLICANT: Robert A. McKay
; APPLICANT: Nicholas M. Dean
; APPLICANT: Brett Monia
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE
; TITLE OF INVENTION: COMPOSITONS AND METHODS FOR THE MODULATION OF JNK
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; MEDIUM TYPE: STORAGE
; COMPUTER: PENTIUM
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,629A
; FILING DATE: August 13, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0215
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: Yes
US-08-910-629A-31

Query Match 0.9%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 28;

```
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 2
US-08-910-629A-42
; Sequence 42, Application US/08910629A
; Patent No. 5877309
;  GENERAL INFORMATION:
;    APPLICANT:  Robert A. McKay
;    APPLICANT:  Nicholas M. Dean
;    APPLICANT:  Brett Monia
;    TITLE OF INVENTION:  ANTISENSE OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR THE MODULATION OF JNK
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Jane Massey Licata
;      STREET:  66 East Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 MB
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  PENTIUM
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,629A
;      FILING DATE:  August 13, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
US-08-910-629A-42

  Query Match             0.9%;  Score 18;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db          1 GACTTTGGCCTGGCCCGG 18


RESULT 3
US-09-209-668-7/c
; Sequence 7, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
;  APPLICANT: Monia, Brett P.
;  APPLICANT: Xu, Xiaoxing S.
;  TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
;  TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
;  FILE REFERENCE: ISPH-0336
;  CURRENT APPLICATION NUMBER: US/09/209,668A
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: antisense sequence
US-09-209-668-7

  Query Match             0.9%;  Score 18;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 4
US-09-287-796-31/c
; Sequence 31, Application US/09287796A
; Patent No. 6133246
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0350
;  CURRENT APPLICATION NUMBER: US/09/287,796A
;  CURRENT FILING DATE: 1999-04-07
;  EARLIER APPLICATION NUMBER: 09/130,616
;  EARLIER FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 31
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-287-796-31

  Query Match             0.9%;  Score 18;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 5
US-09-287-796-42
; Sequence 42, Application US/09287796A
; Patent No. 6133246
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0350
;  CURRENT APPLICATION NUMBER: US/09/287,796A
```

; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: 09/130,616
; EARLIER FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 08/910,629
; EARLIER FILING DATE: 1997-08-03
; NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 42
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-287-796-42

Query Match 0.9%; Score 18; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1300 GACTTTGGCCTGGCCCGG 1317
||||||||||||||||||
Db 1 GACTTTGGCCTGGCCCGG 18

RESULT 6
US-09-130-616-31/c
; Sequence 31, Application US/09130616C
; Patent No. 6221850
; GENERAL INFORMATION:
; APPLICANT: McKay, Robert A.
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Monia, Brett
; APPLICANT: Nero, Pam
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
; FILE REFERENCE: ISPH-0318
; CURRENT APPLICATION NUMBER: US/09/130,616C
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 08/910,629
; EARLIER FILING DATE: 1997-08-03
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 31
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-130-616-31

Query Match 0.9%; Score 18; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1300 GACTTTGGCCTGGCCCGG 1317
||||||||||||||||||
Db 20 GACTTTGGCCTGGCCCGG 3

RESULT 7
US-09-130-616-42
; Sequence 42, Application US/09130616C
; Patent No. 6221850
; GENERAL INFORMATION:
; APPLICANT: McKay, Robert A.
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Monia, Brett
; APPLICANT: Nero, Pam
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
; FILE REFERENCE: ISPH-0318
; CURRENT APPLICATION NUMBER: US/09/130,616C
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 08/910,629
; EARLIER FILING DATE: 1997-08-03
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 42
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-130-616-42

Query Match 0.9%; Score 18; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1300 GACTTTGGCCTGGCCCGG 1317
||||||||||||||||||
Db 1 GACTTTGGCCTGGCCCGG 18

RESULT 8
US-08-730-876-2/c
; Sequence 2, Application US/08730876
; Patent No. 5859314
; GENERAL INFORMATION:
; APPLICANT: HIBBS, Margaret L.;
; APPLICANT: DUNN, Ashley R.;
; APPLICANT: GRAILL, Dianne;
; APPLICANT: HODGSON George;
; APPLICANT: TARLINGTON, David M.;
; APPLICANT: ARMES, Jane
; TITLE OF INVENTION: ANIMALS WITH TARGETED GENE DELETION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/730,876
; FILING DATE: 18-Oct-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/005,578
; FILING DATE: 20-Oct-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5859314man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5369 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-730-876-2

Query Match 0.8%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 916 GGGCAGTTTGGGGAAGT 932
|||||||||||||||||

```
Db         17 GGGCAGTTTGGGGAAGT 1


RESULT 9
US-09-490-692-71/c
; Sequence 71, Application US/09490692
; Patent No. 6180353
; GENERAL INFORMATION:
;  APPLICANT: Nicholas M. Dean
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF DAXX EXPRESSION
;  FILE REFERENCE: RTS-0120
;  CURRENT APPLICATION NUMBER: US/09/490,692
;  CURRENT FILING DATE: 2000-01-24
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 71
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-490-692-71

  Query Match             0.8%;  Score 17;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 TCAGGAGGATGATGAAG 44
              |||||||||||||||||
Db         18 TCAGGAGGATGATGAAG 2


RESULT 10
US-08-222-616-2/c
; Sequence 2, Application US/08222616
; Patent No. 5635177
;  GENERAL INFORMATION:
;    APPLICANT:  Bennett, Brian D.
;    APPLICANT:  Goeddel, David
;    APPLICANT:  Lee, James M.
;    APPLICANT:  Matthews, William
;    APPLICANT:  Tsai, Siao Ping
;    APPLICANT:  Wood, William I.
;    TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST
;    TITLE OF INVENTION:  ANTIBODIES
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,616
;      FILING DATE:  4-APR-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/00586
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/826935
;      FILING DATE:  22-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:  821P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-222-616-2

  Query Match             0.8%;  Score 17;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 11
US-08-446-648-2/c
; Sequence 2, Application US/08446648
; Patent No. 6331302
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  Bennett, Brian D.
;    APPLICANT:  Goeddel, David
;    APPLICANT:  Lee, James M.
;    APPLICANT:  Matthews, William
;    APPLICANT:  Tsai, Siao Ping
;    APPLICANT:  Wood, William I.
;    TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,648
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/222616
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0821P3PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-446-648-2

  Query Match             0.8%;  Score 17;  DB 4;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 12
PCT-US95-04228-2/c
; Sequence 2, Application PC/TUS9504228
;  GENERAL INFORMATION:
;     APPLICANT:  Genentech, Inc.
;     APPLICANT:  Bennett, Brian D.
;     APPLICANT:  Goeddel, David
;     APPLICANT:  Lee, James M.
;     APPLICANT:  Matthews, William
;     APPLICANT:  Tsai, Siao Ping
;     APPLICANT:  Wood, William I.
;     TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
;     NUMBER OF SEQUENCES:  45
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/04228
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/222616
;       FILING DATE:  04-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wendy M. Lee
;       REGISTRATION NUMBER:  00,000
;       REFERENCE/DOCKET NUMBER:  821P3PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1994
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  23 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
PCT-US95-04228-2

  Query Match             0.8%;  Score 17;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 13
US-09-506-073-82/c
; Sequence 82, Application US/09506073
; Patent No. 6410518
; GENERAL INFORMATION:
;  APPLICANT: Monia, Brett P.
;  TITLE OF INVENTION: Antisense Oligonucleotide Modulation of raf Gene Expression
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/506,073
;  CURRENT FILING DATE: 2000-02-18
;  EARLIER APPLICATION NUMBER: US 09/143,214
;  EARLIER FILING DATE: 1998-08-28
;  EARLIER APPLICATION NUMBER: PCT/US98/13961
;  EARLIER FILING DATE: 1998-07-06
;  EARLIER APPLICATION NUMBER: US 08/888,982
;  EARLIER FILING DATE: 1997-07-07
;  EARLIER APPLICATION NUMBER: US 08/756,806
;  EARLIER FILING DATE: 1996-11-26
;  EARLIER APPLICATION NUMBER: PCT/US95/07111
;  EARLIER FILING DATE: 1995-05-31
;  EARLIER APPLICATION NUMBER: US 08/250,856
;  EARLIER FILING DATE: 1994-05-31
;  NUMBER OF SEQ ID NOS: 130
; SEQ ID NO 82
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: antisense sequence
US-09-506-073-82

  Query Match             0.8%;  Score 16;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        155 GAGCGGGCGCCAGGAT 170
              ||||||||||||||||
Db         20 GAGCGGGCGCCAGGAT 5


RESULT 14
US-08-859-998-598
; Sequence 598, Application US/08859998
; Patent No. 5994076
;  GENERAL INFORMATION:
;     APPLICANT:  Chenchik, Alex
;     APPLICANT:  Jokhadze, George
;     APPLICANT:  Bibilashvilli, Robert
;     TITLE OF INVENTION:  METHOD OF ASSAYING DIFFERENTIAL
;     TITLE OF INVENTION:  EXPRESSION
;     NUMBER OF SEQUENCES:  1375
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson, P.C.
;       STREET:  2200 Sand Hill Road, Suite 100
;       CITY:  Menlo Park
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94025
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows95
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/859,998
;       FILING DATE:  21-MAY-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Field, Bret E.
;       REGISTRATION NUMBER:  37,620
;       REFERENCE/DOCKET NUMBER:  09096/002001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-322-5070
;       TELEFAX:  415-854-0875
;   INFORMATION FOR SEQ ID NO:  598:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
```

```
;    FEATURE:
;       OTHER INFORMATION: oligonucleotide primer
US-08-859-998-598

  Query Match              0.8%;  Score 16;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1236 CATCCACCGAGACCTC 1251
              ||||||||||||||||
Db          8 CATCCACCGAGACCTC 23


RESULT 15
US-09-225-928-598
; Sequence 598, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/225,928
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 598:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 24 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION:  oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 598:
US-09-225-928-598

  Query Match              0.8%;  Score 16;  DB 4;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1236 CATCCACCGAGACCTC 1251
              ||||||||||||||||
Db          8 CATCCACCGAGACCTC 23


Search completed: July  4, 2003, 10:39:34
Job time : 100 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 02:17:53 ; Search time 98 Seconds
(without alignments)
6305.650 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	26	1.3	1602	1	US-07-820-011A-1	Sequence 1, Appli
	2	26	1.3	1602	5	PCT-US93-00445-1	Sequence 1, Appli
	3	23	1.1	147	1	US-08-456-647B-25	Sequence 25, Appl
	4	23	1.1	147	2	US-08-237-401A-25	Sequence 25, Appl
	5	23	1.1	170	1	US-08-306-691B-29	Sequence 29, Appl
	6	23	1.1	170	5	PCT-US93-06251-71	Sequence 71, Appl
	7	23	1.1	1611	1	US-07-820-011A-3	Sequence 3, Appli
	8	23	1.1	1611	5	PCT-US93-00445-3	Sequence 3, Appli
	9	22	1.1	1574	3	US-09-173-581-12	Sequence 12, Appl
	10	22	1.1	1574	4	US-09-420-915-12	Sequence 12, Appl
	11	21	1.0	3398	5	PCT-US95-08493-12	Sequence 12, Appl
c	12	20	1.0	579	1	US-08-093-383-2	Sequence 2, Appli
	13	20	1.0	675	1	US-08-707-793A-3	Sequence 3, Appli
	14	20	1.0	675	1	US-08-707-792A-3	Sequence 3, Appli
	15	20	1.0	1548	4	US-09-099-053-1	Sequence 1, Appli
	16	20	1.0	2647	5	PCT-US93-06251-77	Sequence 77, Appl
	17	19	0.9	1392	4	US-09-130-616-171	Sequence 171, App
	18	19	0.9	1523	4	US-09-130-616-172	Sequence 172, App
	19	19	0.9	1619	4	US-09-130-616-173	Sequence 173, App
	20	19	0.9	1769	1	US-08-253-155A-19	Sequence 19, Appl
	21	19	0.9	1780	5	PCT-US94-12913A-17	Sequence 17, Appl
	22	19	0.9	1782	3	US-09-209-668-16	Sequence 16, Appl
	23	19	0.9	1782	4	US-09-130-616-169	Sequence 169, App
c	24	19	0.9	3362	3	US-09-344-519-1	Sequence 1, Appli
c	25	19	0.9	3362	4	US-09-766-387-1	Sequence 1, Appli
	26	19	0.9	3561	1	US-08-097-997A-12	Sequence 12, Appl
	27	19	0.9	3561	3	US-08-665-574C-12	Sequence 12, Appl
	28	19	0.9	3561	4	US-08-946-994-12	Sequence 12, Appl
	29	19	0.9	3671	1	US-08-176-620A-5	Sequence 5, Appli
	30	19	0.9	3671	1	US-08-463-862-5	Sequence 5, Appli
	31	19	0.9	3671	2	US-08-461-985-5	Sequence 5, Appli
	32	19	0.9	3671	2	US-08-458-887-5	Sequence 5, Appli
	33	19	0.9	3671	4	US-08-932-787B-5	Sequence 5, Appli
	34	19	0.9	3671	4	US-08-932-012C-5	Sequence 5, Appli
	35	19	0.9	3671	4	US-08-888-818C-5	Sequence 5, Appli
c	36	18	0.9	20	2	US-08-910-629A-31	Sequence 31, Appl
	37	18	0.9	20	2	US-08-910-629A-42	Sequence 42, Appl
c	38	18	0.9	20	3	US-09-209-668-7	Sequence 7, Appli
c	39	18	0.9	20	3	US-09-287-796-31	Sequence 31, Appl
	40	18	0.9	20	3	US-09-287-796-42	Sequence 42, Appl
c	41	18	0.9	20	4	US-09-130-616-31	Sequence 31, Appl
	42	18	0.9	20	4	US-09-130-616-42	Sequence 42, Appl
	43	18	0.9	179	6	5508199-5	Patent No. 5508199
c	44	18	0.9	479	4	US-09-328-111-432	Sequence 432, App
c	45	18	0.9	572	4	US-09-328-111-87	Sequence 87, Appl

ALIGNMENTS

```
RESULT 1
US-07-820-011A-1
; Sequence 1, Application US/07820011A
; Patent No. 5336615
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Leonard
;    APPLICANT:  Madri, Joseph A.
;    APPLICANT:  Warren, Stephen L.
;    APPLICANT:  Luthringer, Daniel J.
;    TITLE OF INVENTION:  Genetically Engineered
;    TITLE OF INVENTION:  Endothelial Cells Exhibiting Enhanced
;    TITLE OF INVENTION:  Migration
;    TITLE OF INVENTION:  and Plasminogen Activator Activity
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairifield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb storage
;      COMPUTER:  IBM PC XT
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 2.10
;      SOFTWARE:  Displaywrite 3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/820,011A
;      FILING DATE:  19920106
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  LB-101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255 1400
;      TELEFAX:  (203) 254 1101
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1602 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Gallus, gallus
;    PUBLICATION INFORMATION:
;      AUTHORS:  Takeya, Tatsuo
```

```
;      AUTHORS:  Hanafusa, Hidesaburo
;      TITLE:  Structure and Sequence of the
;      TITLE:  Cellular Gene Homologous to the RSV src
;      TITLE:  Gene and the Mechanism for Generating the
;      TITLE:  Transforming Virus
;      JOURNAL:  Cell
;      VOLUME:  32
;      PAGES:  881-890
;      DATE:  March, 1983
US-07-820-011A-1

  Query Match             1.3%;  Score 26;  DB 1;  Length 1602;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1354 GCCAAGTTCCCCATCAAGTGGACAGC 1379
             ||||||||||||||||||||||||||
Db       1264 GCCAAGTTCCCCATCAAGTGGACAGC 1289


RESULT 2
PCT-US93-00445-1
; Sequence 1, Application PC/TUS9300445
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Leonard
;    APPLICANT:  Madri, Joseph A.
;    APPLICANT:  Warren, Stephen L.
;    APPLICANT:  Luthringer, Daniel J.
;    TITLE OF INVENTION:  Genetically Engineered
;    TITLE OF INVENTION:  Endothelial Cells
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 760 Kb storage
;      COMPUTER:  DELL 486/50
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  Displaywrite 3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/00445
;      FILING DATE:  19930105
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/820,011
;      FILING DATE:  06-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-101PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255 1400
;      TELEFAX:  (203) 254 1101
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1602 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Gallus, gallus
;    PUBLICATION INFORMATION:
;      AUTHORS:  Takeya, Tatsuo
;      AUTHORS:  Hanafusa, Hidesaburo
;      TITLE:  Structure and Sequence of the
;      TITLE:  Cellular Gene Homologous to the RSV src
;      TITLE:  Gene and the Mechanism for Generating the
;      TITLE:  Transforming Virus
;      JOURNAL:  Cell
;      VOLUME:  32
;      PAGES:  881-890
;      DATE:  March, 1983
PCT-US93-00445-1

  Query Match             1.3%;  Score 26;  DB 5;  Length 1602;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1354 GCCAAGTTCCCCATCAAGTGGACAGC 1379
             ||||||||||||||||||||||||||
Db       1264 GCCAAGTTCCCCATCAAGTGGACAGC 1289


RESULT 3
US-08-456-647B-25
; Sequence 25, Application US/08456647B
; Patent No. 5811516
;  GENERAL INFORMATION:
;    APPLICANT:  Lemke Ph.D. et al., Greg E.
;    TITLE OF INVENTION:  PROTEIN-TYROSINE KINASE GENES
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,647B
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/237,401
;      FILING DATE:  02-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/884,486
;      FILING DATE:  15-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell Ph.D., John R.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  07251/007002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 678-5070
;      TELEFAX:  (619) 678-5099
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  147 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    IMMEDIATE SOURCE:
;      CLONE:  Tyro-13
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..147
US-08-456-647B-25

  Query Match             1.1%;  Score 23;  DB 1;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 0.094;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1367 TCAAGTGGACAGCTCCTGAAGCC 1389
```

```
                    ||||||||||||||||||||||
Db          95 TCAAGTGGACAGCTCCTGAAGCC 117


RESULT 4
US-08-237-401A-25
; Sequence 25, Application US/08237401A
; Patent No. 5837448
;  GENERAL INFORMATION:
;    APPLICANT:  Lemke Ph.D. et al., Greg E.
;    TITLE OF INVENTION:  PROTEIN-TYROSINE KINASE GENES
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/237,401A
;      FILING DATE:  02-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/884,486
;      FILING DATE:  15-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile Ph.D., Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07251/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 678-5070
;      TELEFAX:  (619) 678-5099
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  147 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    IMMEDIATE SOURCE:
;      CLONE:  Tyro-13
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..147
US-08-237-401A-25

  Query Match             1.1%;  Score 23;  DB 2;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 0.094;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1367 TCAAGTGGACAGCTCCTGAAGCC 1389
                ||||||||||||||||||||||
Db          95 TCAAGTGGACAGCTCCTGAAGCC 117


RESULT 5
US-08-306-691B-29
; Sequence 29, Application US/08306691B
; Patent No. 5734039
;  GENERAL INFORMATION:
;    APPLICANT:  Calabretta, Bruno
;    APPLICANT:  Skorski, Tomasz
;    TITLE OF INVENTION:  ANTISENSE
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seidel, Gonda, Lavorgna & Monaco, P.C.
;      STREET:  Two Penn Center, Suite 1800
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,691B
;      FILING DATE:  September 15, 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  8321-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5734039e
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  170 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-306-691B-29

  Query Match             1.1%;  Score 23;  DB 1;  Length 170;
  Best Local Similarity   100.0%;  Pred. No. 0.095;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 GATGCCTGGGAGATCCCTCGGGA 884
                ||||||||||||||||||||||
Db         87 GATGCCTGGGAGATCCCTCGGGA 109


RESULT 6
PCT-US93-06251-71
; Sequence 71, Application PC/TUS9306251
;  GENERAL INFORMATION:
;    APPLICANT:  Wickstrom, Eric and Rife, Jason P.
;    TITLE OF INVENTION:  Trivalent Synthesis of Oligonucleotides Containing
;    TITLE OF INVENTION:  Stereospecific Alkylphosphonates and Arylphosphonates
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/06251
;      FILING DATE:  19930630
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  516-742-4343
;      TELEFAX:  516-742-4366
```

```
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  71:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  170 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US93-06251-71

  Query Match             1.1%;  Score 23;  DB 5;  Length 170;
  Best Local Similarity   100.0%;  Pred. No. 0.095;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 GATGCCTGGGAGATCCCTCGGGA 884
              |||||||||||||||||||||||
Db         87 GATGCCTGGGAGATCCCTCGGGA 109


RESULT 7
US-07-820-011A-3
; Sequence 3, Application US/07820011A
; Patent No. 5336615
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Leonard
;    APPLICANT:  Madri, Joseph A.
;    APPLICANT:  Warren, Stephen L.
;    APPLICANT:  Luthringer, Daniel J.
;    TITLE OF INVENTION:  Genetically Engineered
;    TITLE OF INVENTION:  Endothelial Cells Exhibiting Enhanced
;    TITLE OF INVENTION:  Migration
;    TITLE OF INVENTION:  and Plasminogen Activator Activity
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairifield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb storage
;      COMPUTER:  IBM PC XT
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 2.10
;      SOFTWARE:  Displaywrite 3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/820,011A
;      FILING DATE:  19920106
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  LB-101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255 1400
;      TELEFAX:  (203) 254 1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1611
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  Chromosome 20
;    PUBLICATION INFORMATION:
;      AUTHORS:  Anderson, Stephen K.
;      AUTHORS:  Gibbs, Carol P.
;      AUTHORS:  Tanaka, Akio
;      AUTHORS:  Kung, Hsing-Jien
;      AUTHORS:  Fujita, Donald J.
;      TITLE:  Human Cellular src Gene:
;      TITLE:  Nucleotide Sequence and Derived Amino
;      TITLE:  Acid Sequence of the Region Coding for
;      TITLE:  the Carboxy-Terminal Two-Thirds of
;      TITLE:  pp60c-src
;      JOURNAL:  Molecular and Cellular Biology
;      VOLUME:  5
;      ISSUE:  5
;      PAGES:  1122-1129
;      DATE:  May, 1985
;    PUBLICATION INFORMATION:
;      AUTHORS:  Tanaka, Akio
;      AUTHORS:  Gibbs, Carol P.
;      AUTHORS:  Arthur, Richard R.
;      AUTHORS:  Anderson, Stephen K.
;      AUTHORS:  Kung, Hsing-Jien
;      AUTHORS:  Fujita, Donald J.
;      TITLE:  DNA Sequence Encoding the
;      TITLE:  Amino-Terminal Region of the Human c-src
;      TITLE:  Protein: Implications of Sequence
;      TITLE:  Divergence among src-Type Kinase
;      TITLE:  Oncogenes
;      JOURNAL:  Molecular and Cellular Biology
;      VOLUME:  7
;      ISSUE:  5
;      PAGES:  1978-1983
;      DATE:  May, 1987
US-07-820-011A-3

  Query Match             1.1%;  Score 23;  DB 1;  Length 1611;
  Best Local Similarity   100.0%;  Pred. No. 0.099;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 GATGCCTGGGAGATCCCTCGGGA 884
              |||||||||||||||||||||||
Db        781 GATGCCTGGGAGATCCCTCGGGA 803


RESULT 8
PCT-US93-00445-3
; Sequence 3, Application PC/TUS9300445
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Leonard
;    APPLICANT:  Madri, Joseph A.
;    APPLICANT:  Warren, Stephen L.
;    APPLICANT:  Luthringer, Daniel J.
;    TITLE OF INVENTION:  Genetically Engineered
;    TITLE OF INVENTION:  Endothelial Cells
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 760 Kb storage
;      COMPUTER:  DELL 486/50
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  Displaywrite 3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/00445
;      FILING DATE:  19930105
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/820,011
;      FILING DATE:  06-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
```

```
;      REFERENCE/DOCKET NUMBER:  ALX-101PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255 1400
;      TELEFAX:  (203) 254 1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1611
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  Chromosome 20
;    PUBLICATION INFORMATION:
;      AUTHORS:  Anderson, Stephen K.
;      AUTHORS:  Gibbs, Carol P.
;      AUTHORS:  Tanaka, Akio
;      AUTHORS:  Kung, Hsing-Jien
;      AUTHORS:  Fujita, Donald J.
;      TITLE:  Human Cellular src Gene:
;      TITLE:  Nucleotide Sequence and Derived Amino
;      TITLE:  Acid Sequence of the Region Coding for
;      TITLE:  the Carboxy-Terminal Two-Thirds of
;      TITLE:  pp60c-src
;      JOURNAL:  Molecular and Cellular Biology
;      VOLUME:  5
;      ISSUE:  5
;      PAGES:  1122-1129
;      DATE:  May, 1985
;    PUBLICATION INFORMATION:
;      AUTHORS:  Tanaka, Akio
;      AUTHORS:  Gibbs, Carol P.
;      AUTHORS:  Arthur, Richard R.
;      AUTHORS:  Anderson, Stephen K.
;      AUTHORS:  Kung, Hsing-Jien
;      AUTHORS:  Fujita, Donald J.
;      TITLE:  DNA Sequence Encoding the
;      TITLE:  Amino-Terminal Region of the Human c-src
;      TITLE:  Protein: Implications of Sequence
;      TITLE:  Divergence among src-Type Kinase
;      TITLE:  Oncogenes
;      JOURNAL:  Molecular and Cellular Biology
;      VOLUME:  7
;      ISSUE:  5
;      PAGES:  1978-1983
;      DATE:  May, 1987
PCT-US93-00445-3

  Query Match             1.1%;  Score 23;  DB 5;  Length 1611;
  Best Local Similarity   100.0%;  Pred. No. 0.099;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 GATGCCTGGGAGATCCCTCGGGA 884
              |||||||||||||||||||||||
Db        781 GATGCCTGGGAGATCCCTCGGGA 803


RESULT 9
US-09-173-581-12
; Sequence 12, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/173,581A
;  CURRENT FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 12
;   LENGTH: 1574
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 507669
US-09-173-581-12

  Query Match             1.1%;  Score 22;  DB 3;  Length 1574;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1082 AGGAGCCCATCTACATCATCAC 1103
              ||||||||||||||||||||||
Db        532 AGGAGCCCATCTACATCATCAC 553


RESULT 10
US-09-420-915-12
; Sequence 12, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/420,915
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/173,581
;  EARLIER FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 12
;   LENGTH: 1574
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 507669
US-09-420-915-12

  Query Match             1.1%;  Score 22;  DB 4;  Length 1574;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1082 AGGAGCCCATCTACATCATCAC 1103
              ||||||||||||||||||||||
Db        532 AGGAGCCCATCTACATCATCAC 553


RESULT 11
PCT-US95-08493-12
; Sequence 12, Application PC/TUS9508493
;  GENERAL INFORMATION:
;    APPLICANT:  Wood, Clive
;    APPLICANT:  Caruso, Anthony
;    TITLE OF INVENTION:  Novel mlk Receptor Tyrosine Kinases
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  LEGAL AFFAIRS
;       STREET:  87 CambridgePark Drive
;       CITY:  Cambridge
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02140
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/08493
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brown, Scott A
;       REGISTRATION NUMBER:  32,724
;       REFERENCE/DOCKET NUMBER:  GI5234A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 498-8224
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3398 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  121..2961
PCT-US95-08493-12

  Query Match             1.0%;  Score 21;  DB 5;  Length 3398;
  Best Local Similarity   100.0%;  Pred. No. 1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1290 TAAGATTGCTGACTTTGGCCT 1310
              |||||||||||||||||||||
Db       2568 TAAGATTGCTGACTTTGGCCT 2588


RESULT 12
US-08-093-383-2/c
; Sequence 2, Application US/08093383
; Patent No. 5489529
;  GENERAL INFORMATION:
;     APPLICANT:  DeBoer, Herman A.
;     APPLICANT:  Heyneker, Herbert L.
;     APPLICANT:  Seeburg, Peter H.
;     TITLE OF INVENTION:  DNA for Expression of Bovine Growth Hormone
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/093,383
;       FILING DATE:  14-JUL-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/619827
;       FILING DATE:  28-NOV-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/198824
;       FILING DATE:  05-APR-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  06/632361
;       FILING DATE:  19-JUL-1984
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  06/303687
;       FILING DATE:  18-SEP-1981
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Johnston, Sean A.
;       REGISTRATION NUMBER:  P35,910
;       REFERENCE/DOCKET NUMBER:  46C4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-3562
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  579 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-093-383-2

  Query Match             1.0%;  Score 20;  DB 1;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         56 GATGAAGACGATGACGACGA 75
              ||||||||||||||||||||
Db         67 GATGAAGACGATGACGACGA 48


RESULT 13
US-08-707-793A-3
; Sequence 3, Application US/08707793A
; Patent No. 5776696
;  GENERAL INFORMATION:
;     APPLICANT:  SALOWE, SCOTT P.
;     TITLE OF INVENTION:  A HIGH THROUGHPUT ASSAY USING
;     TITLE OF INVENTION:  FUSION PROTEINS
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Merck & Co., Inc.
;       STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;       CITY:  Rahway
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07065-0900
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/707,793A
;       FILING DATE:  04-SEP-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Camara, Valerie J
;       REGISTRATION NUMBER:  35,090
;       REFERENCE/DOCKET NUMBER:  19494
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  908-594-3902
;       TELEFAX:  908-594-4720
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  675 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-707-793A-3

  Query Match             1.0%;  Score 20;  DB 1;  Length 675;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        688 GTGAAACATTACAAGATCCG 707
              ||||||||||||||||||||
Db        526 GTGAAACATTACAAGATCCG 545


RESULT 14
US-08-707-792A-3
; Sequence 3, Application US/08707792A
; Patent No. 5783398
;  GENERAL INFORMATION:
;    APPLICANT:  MARCY, ALICE
;    APPLICANT:  SALOWE, SCOTT P.
;    APPLICANT:  WISNIEWSKI, DOUGLAS
;    TITLE OF INVENTION:  A HIGH THROUGHPUT ASSAY USING
;    TITLE OF INVENTION:  FUSION PROTEINS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/707,792A
;      FILING DATE:  04-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Camara, Valerie J
;      REGISTRATION NUMBER:  35,090
;      REFERENCE/DOCKET NUMBER:  19524
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-3902
;      TELEFAX:  908-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  675 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-707-792A-3

  Query Match             1.0%;  Score 20;  DB 1;  Length 675;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        688 GTGAAACATTACAAGATCCG 707
              ||||||||||||||||||||
Db        526 GTGAAACATTACAAGATCCG 545


RESULT 15
US-09-099-053-1
; Sequence 1, Application US/09099053
; Patent No. 6388063
;  GENERAL INFORMATION:
;    APPLICANT:  Greg Plowman
;    APPLICANT:  Susan Onrust
;    APPLICANT:  David Markby
;    APPLICANT:  Sara Courtneidge
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF
;    TITLE OF INVENTION:  SAD RELATED DISORDERS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/099,053
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/049,914
;      FILING DATE:  June 18, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  235/121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1548 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-099-053-1

  Query Match             1.0%;  Score 20;  DB 4;  Length 1548;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1414 AAGTCAGACGTCTGGTCCTT 1433
              ||||||||||||||||||||
Db       1264 AAGTCAGACGTCTGGTCCTT 1283


Search completed: July  4, 2003, 04:49:24
Job time : 127 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 04:49:28 ; Search time 5237 Seconds
(without alignments)
11197.662 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	26	1.3	72	14	ALRSRCB	J02351 Rous sarcom
	2	25	1.2	51	6	AX165171	AX165171 Sequence
	3	18	0.9	19	6	AX129247	AX129247 Sequence
c	4	18	0.9	20	6	AR110470	AR110470 Sequence
c	5	18	0.9	20	6	AR116450	AR116450 Sequence
	6	18	0.9	20	6	AR116461	AR116461 Sequence
c	7	18	0.9	20	6	AX104119	AX104119 Sequence
c	8	18	0.9	20	6	AX164692	AX164692 Sequence
c	9	18	0.9	20	6	AX355435	AX355435 Sequence
c	10	18	0.9	48	6	AX427069	AX427069 Sequence
c	11	18	0.9	51	6	AX427068	AX427068 Sequence
	12	17	0.8	19	6	AX129246	AX129246 Sequence
c	13	17	0.8	20	6	AR029423	AR029423 Sequence
c	14	17	0.8	20	6	AR126642	AR126642 Sequence
c	15	17	0.8	21	6	AX201544	AX201544 Sequence
c	16	17	0.8	23	6	I44506	I44506 Sequence 2
c	17	17	0.8	57	6	AX179479	AX179479 Sequence
c	18	17	0.8	63	9	AF339072	AF339072 Cheirogal
c	19	17	0.8	63	9	AF339077	AF339077 Pan trogl
	20	17	0.8	71	4	AF055530	AF055530 Didelphis
c	21	16	0.8	22	6	AX465576	AX465576 Sequence
	22	16	0.8	24	6	AR090478	AR090478 Sequence
	23	16	0.8	24	6	AR197513	AR197513 Sequence
c	24	16	0.8	64	6	A67729	A67729 Sequence 59
	25	16	0.8	71	9	HSU38ASNR	X97582 H.sapiens s
c	26	15	0.7	18	6	AR190730	AR190730 Sequence
	27	15	0.7	19	6	I77125	I77125 Sequence 11
	28	15	0.7	31	6	AR069592	AR069592 Sequence
	29	15	0.7	31	6	AX249143	AX249143 Sequence
	30	15	0.7	31	6	AX249144	AX249144 Sequence
	31	15	0.7	31	6	AX249145	AX249145 Sequence
	32	15	0.7	31	6	AX249146	AX249146 Sequence
	33	15	0.7	31	6	AX249147	AX249147 Sequence
	34	15	0.7	36	6	AX069497	AX069497 Sequence
c	35	15	0.7	36	6	AX069498	AX069498 Sequence
	36	15	0.7	43	6	AX141093	AX141093 Sequence
	37	15	0.7	43	6	AX146963	AX146963 Sequence
c	38	15	0.7	44	6	AX473094	AX473094 Sequence
c	39	15	0.7	45	6	AR028569	AR028569 Sequence
	40	15	0.7	48	6	A18448	A18448 oligonucleo
	41	15	0.7	51	6	AX159179	AX159179 Sequence
	42	15	0.7	51	6	AX159180	AX159180 Sequence
c	43	15	0.7	51	6	AX199202	AX199202 Sequence
	44	15	0.7	60	6	AX455886	AX455886 Sequence
c	45	15	0.7	65	6	AX482877	AX482877 Sequence

ALIGNMENTS

RESULT 1
ALRSRCB
LOCUS ALRSRCB 72 bp ss-RNA linear VRL 28-APR-1993
DEFINITION Rous sarcoma virus (RSV) src gene, partial.
ACCESSION J02351
VERSION J02351.1 GI:210266
KEYWORDS c-myc proto-oncogene; kinase; protein kinase; src oncogene.
SOURCE Rous sarcoma virus (Prague A strain) DNA, clones pCH1,pCH7 & pCH20.
ORGANISM Rous sarcoma virus
Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
REFERENCE 1 (bases 1 to 72)
AUTHORS Bryant,D. and Parsons,J.T.
TITLE Site-directed point mutation in the src gene oF rous sarcoma virus results in an inactive src gene product
JOURNAL J. Virol. 45 (3), 1211-1216 (1983)

```
  MEDLINE   83164366
   PUBMED   6300458
COMMENT     A 'g' to 'a' transition at base 55 results in the incorporation of
            Thr instead of Ala at amino acid 433. This change decreases the
            protein kinase activity of the product and abolishes the pp60-src
            mediated cellular transformation activity.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Rous sarcoma virus"
                     /db_xref="taxon:11886"
     CDS             <1. .>72
                     /note="v-src protein"
                     /codon_start=1
                     /protein_id="AAA42584.1"
                     /db_xref="GI:210267"
                     /translation="EYTARQGAKFPIKWTAPEAALYGR"
BASE COUNT       17 a     25 c     21 g      9 t
ORIGIN      52 bp upstream of BglI site.

  Query Match             1.3%;  Score 26;  DB 14;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1354 GCCAAGTTCCCCATCAAGTGGACAGC 1379
              ||||||||||||||||||||||||||
Db         22 GCCAAGTTCCCCATCAAGTGGACAGC 47


RESULT 2
AX165171
LOCUS       AX165171                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 366 from Patent WO0138586.
ACCESSION   AX165171
VERSION     AX165171.1  GI:14546000
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 366 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg42665067"
BASE COUNT       10 a     17 c     13 g     11 t
ORIGIN

  Query Match             1.2%;  Score 25;  DB 6;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 0.045;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        722 GGGGCTTCTACATATCCCCCCGAAG 746
              |||||||||||||||||||||||||
Db          1 GGGGCTTCTACATATCCCCCCGAAG 25


RESULT 3
AX129247
LOCUS       AX129247                 19 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 465 from Patent WO0130362.
ACCESSION   AX129247
VERSION     AX129247.1  GI:14135552
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Tritz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative skin and eye
            diseases
  JOURNAL   Patent: WO 0130362-A 465 03-MAY-2001;
            IMMUSOL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Cdk4 ribozyme binding site"
BASE COUNT        1 a      6 c      7 g      5 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1297 GCTGACTTTGGCCTGGCC 1314
              ||||||||||||||||||
Db          2 GCTGACTTTGGCCTGGCC 19


RESULT 4
AR110470/c
LOCUS       AR110470                 20 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 7 from patent US 6114517.
ACCESSION   AR110470
VERSION     AR110470.1  GI:12826746
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Monia,B.P. and Xu,X.S.
  TITLE     Methods of modulating tumor necrosis factor .alpha.-induced
            expression of cell adhesion molecules
  JOURNAL   Patent: US 6114517-A 7 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      7 c      7 g      2 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 5
AR116450/c
LOCUS       AR116450                 20 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 31 from patent US 6133246.
ACCESSION   AR116450
VERSION     AR116450.1  GI:14096772
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   McKay,R., Dean,N., Monia,B.P., Nero,P.S. and Gaarde,W.A.
  TITLE     Antisense oligonucleotide compositions and methods for the
            modulation of JNK proteins
  JOURNAL   Patent: US 6133246-A 31 17-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
```

```
BASE COUNT        4 a      7 c      7 g      2 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 6
AR116461
LOCUS       AR116461                 20 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 42 from patent US 6133246.
ACCESSION   AR116461
VERSION     AR116461.1  GI:14096783
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   McKay,R., Dean,N., Monia,B.P., Nero,P.S. and Gaarde,W.A.
  TITLE     Antisense oligonucleotide compositions and methods for the
            modulation of JNK proteins
  JOURNAL   Patent: US 6133246-A 42 17-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        2 a      7 c      7 g      4 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db          1 GACTTTGGCCTGGCCCGG 18


RESULT 7
AX104119/c
LOCUS       AX104119                 20 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 311 from Patent WO0122972.
ACCESSION   AX104119
VERSION     AX104119.1  GI:13920316
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 311 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        4 a      7 c      7 g      2 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 8
AX164692/c
LOCUS       AX164692                 20 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 2 from Patent WO0134792.
ACCESSION   AX164692
VERSION     AX164692.1  GI:14545586
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Potapova,O., Gorospe,M. and Holbrook,N.J.
  TITLE     Compositions and methods for the diminution or elimination of
            various cancers
  JOURNAL   Patent: WO 0134792-A 2 17-MAY-2001;
            THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
BASE COUNT        4 a      7 c      7 g      2 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 9
AX355435/c
LOCUS       AX355435                 20 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 463 from Patent WO0197843.
ACCESSION   AX355435
VERSION     AX355435.1  GI:18620103
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Weiner,G. and Hartmann,G.
  TITLE     Methods for enhancing antibody-induced cell lysis and treating
            cancer
  JOURNAL   Patent: WO 0197843-A 463 27-DEC-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide-phosphorothioate
                     backbone"
BASE COUNT        4 a      7 c      7 g      2 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 10
AX427069/c
LOCUS       AX427069                 48 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 33 from Patent WO0196604.
ACCESSION   AX427069
```

```
VERSION     AX427069.1  GI:21530452
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bee,G., Kohne,D.E., Korb,L., Peterson,T. and Yguerabide,J.
  TITLE     Assay for genetic polymorphisms using scattered light detectable
            labels
  JOURNAL   Patent: WO 0196604-A 33 20-DEC-2001;
            Genicon Sciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Exemplary probe for CYP2D6 allele detection"
BASE COUNT        9 a      20 c      10 g       9 t
ORIGIN

  Query Match           0.9%;  Score 18;  DB 6;  Length 48;
  Best Local Similarity   100.0%;  Pred. No. 5e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1003 GAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||
Db         45 GAGGCCTTCCTGGCAGAG 28


RESULT 11
AX427068/c
LOCUS       AX427068                  51 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 32 from Patent WO0196604.
ACCESSION   AX427068
VERSION     AX427068.1  GI:21530451
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bee,G., Kohne,D.E., Korb,L., Peterson,T. and Yguerabide,J.
  TITLE     Assay for genetic polymorphisms using scattered light detectable
            labels
  JOURNAL   Patent: WO 0196604-A 32 20-DEC-2001;
            Genicon Sciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Exemplary probe for CYP2D6 allele detection"
BASE COUNT        9 a      21 c      10 g      11 t
ORIGIN

  Query Match           0.9%;  Score 18;  DB 6;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 5e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1003 GAGGCCTTCCTGGCAGAG 1020
              ||||||||||||||||||
Db         48 GAGGCCTTCCTGGCAGAG 31


RESULT 12
AX129246
LOCUS       AX129246                  19 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 464 from Patent WO0130362.
ACCESSION   AX129246
VERSION     AX129246.1  GI:14135551
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Tritz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative skin and eye
            diseases
  JOURNAL   Patent: WO 0130362-A 464 03-MAY-2001;
            IMMUSOL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Cdk4 ribozyme binding site"
BASE COUNT        1 a      5 c      7 g      6 t
ORIGIN

  Query Match           0.8%;  Score 17;  DB 6;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1297 GCTGACTTTGGCCTGGC 1313
              |||||||||||||||||
Db          3 GCTGACTTTGGCCTGGC 19


RESULT 13
AR029423/c
LOCUS       AR029423                  20 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5859314.
ACCESSION   AR029423
VERSION     AR029423.1  GI:5941396
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Hibbs,M.L., Dunn,A.R., Graill,D., Hodgson,G., Tarlington,D.M. and
            Armes,J.
  TITLE     Mice with targeted tyrosine kinase, lyn, disruption
  JOURNAL   Patent: US 5859314-A 2 12-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      10 c      2 g       4 t
ORIGIN

  Query Match           0.8%;  Score 17;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        916 GGGCAGTTTGGGGAAGT 932
              |||||||||||||||||
Db         17 GGGCAGTTTGGGGAAGT 1


RESULT 14
AR126642/c
LOCUS       AR126642                  20 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 71 from patent US 6180353.
ACCESSION   AR126642
VERSION     AR126642.1  GI:14113235
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dean,N.M. and Cowsert,L.M.
  TITLE     Antisense modulation of daxx expression
  JOURNAL   Patent: US 6180353-A 71 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        3 a      9 c      1 g       7 t
ORIGIN

  Query Match           0.8%;  Score 17;  DB 6;  Length 20;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 TCAGGAGGATGATGAAG 44
              |||||||||||||||||
Db         18 TCAGGAGGATGATGAAG 2


RESULT 15
AX201544/c
LOCUS       AX201544                 21 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 223 from Patent WO0153486.
ACCESSION   AX201544
VERSION     AX201544.1  GI:15391386
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
            Stone,D.M., Watanabe,C.K. and Wood,W.I.
  TITLE     Compositions and methods for the treatment of tumour
  JOURNAL   Patent: WO 0153486-A 223 26-JUL-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide Probe."
BASE COUNT        7 a      8 c      3 g      3 t
ORIGIN

  Query Match             0.8%;  Score 17;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        907 CTTGGAGCTGGGCAGTT 923
              |||||||||||||||||
Db         19 CTTGGAGCTGGGCAGTT 3


Search completed: July  4, 2003, 08:32:26
Job time : 5240 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 06:11:39 ; Search time 2904 Seconds
(without alignments)
11237.571 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18	0.9	93	10	AW238943	AW238943 xb29h03.y
c	2	17	0.8	52	12	BF636617	BF636617 NF091E04D
	3	16	0.8	53	13	BJ048171	BJ048171 BJ048171
	4	16	0.8	56	17	AZ801785	AZ801785 2M0060E22
	5	16	0.8	64	9	AA117806	AA117806 mo65a02.r
	6	16	0.8	66	9	AI906791	AI906791 IL-BT125-
c	7	16	0.8	66	9	AI906801	AI906801 RC-BT125-
	8	16	0.8	66	9	AI906818	AI906818 RC-BT125-
c	9	16	0.8	67	9	AA235459	AA235459 zt31a12.s
c	10	16	0.8	68	10	AV832681	AV832681 AV832681
c	11	16	0.8	72	17	B35762	B35762 HS-1030-B1-
	12	16	0.8	73	17	BH801576	BH801576 1008117H1
	13	16	0.8	79	13	BJ034007	BJ034007 BJ034007
c	14	16	0.8	82	9	AA761095	AA761095 ny13h05.s
c	15	16	0.8	88	9	AA620617	AA620617 af84b06.s
c	16	16	0.8	97	14	H61808	H61808 yu41c09.s1
c	17	15	0.7	33	10	AW059793	AW059793 LE5g11.yg
	18	15	0.7	36	14	D12082	D12082 HUM0S16A04
	19	15	0.7	44	14	D19127	D19127 MUSGS01342
	20	15	0.7	44	17	BH643552	BH643552 1008058F0
c	21	15	0.7	54	17	B33982	B33982 HS-1023-B2-
	22	15	0.7	55	14	R54182	R54182 yg98b12.r1
c	23	15	0.7	55	17	BH415966	BH415966 1007045F1
	24	15	0.7	58	13	BM307600	BM307600 sak31d09.
c	25	15	0.7	58	14	C01969	C01969 HUMGS000398
c	26	15	0.7	59	13	BM018730	BM018730 603646626
c	27	15	0.7	61	14	C02010	C02010 HUMGS000453
c	28	15	0.7	62	14	H25116	H25116 yl43b09.s1
	29	15	0.7	62	17	TA156A09P	AL467282 T. brucei
c	30	15	0.7	63	17	AZ307924	AZ307924 1M0010E24
c	31	15	0.7	64	10	BE636462	BE636462 SWOvL2CAS
c	32	15	0.7	64	17	AZ960593	AZ960593 2M0228N10
	33	15	0.7	67	9	AL644576	AL644576 AL644576
	34	15	0.7	70	9	AI215407	AI215407 qh07f12.x
	35	15	0.7	74	17	CNS030KY	AL222379 Tetraodon
	36	15	0.7	75	14	F33772	F33772 HSPD27429 H
	37	15	0.7	75	17	AZ467917	AZ467917 1M0279E08
c	38	15	0.7	76	9	AA826784	AA826784 nr89f05.s
	39	15	0.7	77	9	AA743582	AA743582 ny29c01.s
c	40	15	0.7	80	9	AA930335	AA930335 vs59c04.r
	41	15	0.7	81	9	AA585407	AA585407 PTH327A H
c	42	15	0.7	81	17	AF179956	AF179956 AF179956
	43	15	0.7	84	17	B40784	B40784 HS-1052-B1-
	44	15	0.7	84	17	CNS03YXT	AL266906 Tetraodon
	45	15	0.7	85	10	AV533467	AV533467 AV533467

ALIGNMENTS

RESULT 1
AW238943
LOCUS AW238943 93 bp mRNA linear EST 13-DEC-1999
DEFINITION xb29h03.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2577749 5' similar to gb:X12597 HIGH MOBILITY GROUP PROTEIN HMG1 (HUMAN);contains element THR repetitive element ;, mRNA sequence.
ACCESSION AW238943
VERSION AW238943.1 GI:6571333
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: xb29h03.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40RP from Gibco

```
          High quality sequence stop: 86.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2577749"
                     /clone_lib="NCI_CGAP_Lu31"
                     /sex="male"
                     /dev_stage="fetal, 14 wk post-conception"
                     /lab_host="DH10B"
                     /note="Organ: lung, cell line; Vector: pCMV-SPORT6;
                     Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally, no
                     5' adaptor.  Primer: Oligo dT.   Full-length library
                     constructed by Life Technologies."
BASE COUNT       46 a      8 c     30 g      9 t
ORIGIN

  Query Match             0.9%;  Score 18;  DB 10;  Length 93;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 GGAAGATGAGGAAGATGA 27
              ||||||||||||||||||
Db         75 GGAAGATGAGGAAGATGA 92


RESULT 2
BF636617/c
LOCUS       BF636617                 52 bp    mRNA    linear   EST 19-DEC-2000
DEFINITION  NF091E04DT1F1034 Drought Medicago truncatula cDNA clone NF091E04DT
            5', mRNA sequence.
ACCESSION   BF636617
VERSION     BF636617.1  GI:11900775
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Torrez-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula drought library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Insert Length: 52   Std Error: 0.00
            Plate: 091  row: E  column: 04
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF091E04DT"
                     /clone_lib="Drought"
                     /tissue_type="Plantlets"
                     /dev_stage="Pooled timepoints"
                     /note="Vector: Lambda Zap; Contains a mixture of entire
                     plantlets harvested in a series of days-post-watering
                     timepoints. "
BASE COUNT       12 a      5 c      2 g     22 t     11 others
ORIGIN

  Query Match             0.8%;  Score 17;  DB 12;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1966 AAAATAGTGAAATGAAT 1982
              |||||||||||||||||
Db         39 AAAATAGTGAAATGAAT 23


RESULT 3
BJ048171
LOCUS       BJ048171                 53 bp    mRNA    linear   EST 07-DEC-2001
DEFINITION  BJ048171 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL018o16 3', mRNA sequence.
ACCESSION   BJ048171
VERSION     BJ048171.1  GI:17406228
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL018o16"
                     /clone_lib="NIBB Mochii normalized Xenopus neurula
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 15"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       16 a      9 c     13 g     11 t      4 others
ORIGIN

  Query Match             0.8%;  Score 16;  DB 13;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 8.2e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        845 AGAAGCCTTGGGAGAA 860
              ||||||||||||||||
Db         20 AGAAGCCTTGGGAGAA 35


RESULT 4
AZ801785
LOCUS       AZ801785                 56 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0060E22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0060E22 F, DNA sequence.
ACCESSION   AZ801785
VERSION     AZ801785.1  GI:12954108
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
```

```
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0060  row: E  column: 22
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0060E22"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a      3 c     24 g     11 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 17;  Length 56;
  Best Local Similarity   100.0%;  Pred. No. 8.3e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 GTGAAGAGGGAGATGA 60
              ||||||||||||||||
Db         12 GTGAAGAGGGAGATGA 27


RESULT 5
AA117806
LOCUS       AA117806                 64 bp    mRNA    linear   EST 15-FEB-1997
DEFINITION  mo65a02.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
            IMAGE:558410 5', mRNA sequence.
ACCESSION   AA117806
VERSION     AA117806.1  GI:1672822
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:339202
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 52.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Mus musculus"
                     /strain="NIH Swiss"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:558410"
                     /clone_lib="Stratagene mouse heart (#937316)"
                     /sex="pooled"
                     /tissue_type="heart"
                     /dev_stage="13 day embryos"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: heart; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       26 a      3 c     29 g      6 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 8.5e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         41 GAAGGTGAAGAGGGAG 56
              ||||||||||||||||
Db         27 GAAGGTGAAGAGGGAG 42


RESULT 6
AI906791
LOCUS       AI906791                 66 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  IL-BT125-090299-004 BT125 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI906791
VERSION     AI906791.1  GI:6497199
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
```

```
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT125-004.html
            &t3=090299&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .66
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT125"
                      /sex="female"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       18 a     12 c     21 g     15 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1575 GATGCGCTGCTGGAAA 1590
              ||||||||||||||||
Db         45 GATGCGCTGCTGGAAA 60


RESULT 7
AI906801/c
LOCUS       AI906801                 66 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  RC-BT125-030399-008 BT125 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI906801
VERSION     AI906801.1  GI:6497209
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=RC&t2=RC-BT125-008.html
            &t3=030399&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .66
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT125"
                      /sex="female"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       15 a     21 c     12 g     18 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1575 GATGCGCTGCTGGAAA 1590
              ||||||||||||||||
Db         22 GATGCGCTGCTGGAAA 7


RESULT 8
AI906818
LOCUS       AI906818                 66 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  RC-BT125-040399-020 BT125 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI906818
VERSION     AI906818.1  GI:6497226
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=RC&t2=RC-BT125-020.html
            &t3=040399&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .66
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT125"
                      /sex="female"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       18 a     12 c     21 g     15 t
ORIGIN
```

```
  Query Match             0.8%;  Score 16;  DB 9;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1575 GATGCGCTGCTGGAAA 1590
              ||||||||||||||||
Db         45 GATGCGCTGCTGGAAA 60


RESULT 9
AA235459/c
LOCUS       AA235459                  67 bp    mRNA    linear   EST 08-AUG-1997
DEFINITION  zt31a12.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:723934 3' similar to SW:KR4_HUMAN P10075 HKR4 PROTEIN ;, mRNA
            sequence.
ACCESSION   AA235459
VERSION     AA235459.1  GI:1860110
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Insert Length: 362   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:723934"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT        6 a     30 c     19 g     12 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         96 GGGCTGCCGAGCTGGG 111
              ||||||||||||||||
Db         63 GGGCTGCCGAGCTGGG 48


RESULT 10
AV832681/c
LOCUS       AV832681                  68 bp    mRNA    linear   EST 22-JUN-2001
DEFINITION  AV832681 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare top three leaves adult, heading stage Hordeum vulgare
            subsp. vulgare cDNA clone baa10f10, mRNA sequence.
ACCESSION   AV832681
VERSION     AV832681.1  GI:14524770
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
            submission;
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="baa10f10"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare top three leaves adult, heading
                     stage"
                     /tissue_type="top three leaves"
                     /dev_stage="adult, heading stage"
BASE COUNT       15 a     22 c     18 g     12 t      1 others
ORIGIN

  Query Match             0.8%;  Score 16;  DB 10;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1800 CAGCCACAGTTTCCTC 1815
              ||||||||||||||||
Db         61 CAGCCACAGTTTCCTC 46


RESULT 11
B35762/c
LOCUS       B35762                    72 bp    DNA     linear   GSS 17-OCT-1997
DEFINITION  HS-1030-B1-B02-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT810 Col=3 Row=D, DNA sequence.
ACCESSION   B35762
VERSION     B35762.1  GI:2535131
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
```

```
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT810  row: D  column: 3
            Class: BAC ends
            High quality sequence stop: 72.
FEATURES              Location/Qualifiers
     source           1. .72
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="Plate=CT810 Col=3 Row=D"
                      /clone_lib="CIT Human Genomic Sperm Library C"
                      /sex="M"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
BASE COUNT       13 a     20 c     17 g     21 t      1 others
ORIGIN

  Query Match             0.8%;  Score 16;  DB 17;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 8.7e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1020 GGCCAACGTGATGAAA 1035
              ||||||||||||||||
Db         60 GGCCAACGTGATGAAA 45


RESULT 12
BH801576
LOCUS       BH801576                  73 bp    DNA     linear   GSS 25-APR-2002
DEFINITION  1008117H12.2EL_y1 1008 - RescueMu Grid I Zea mays genomic, DNA
            sequence.
ACCESSION   BH801576
VERSION     BH801576.1  GI:20314787
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008117  row: 20
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Zea mays"
                      /cultivar="mixed background W23/A188/B73"
                      /db_xref="taxon:4577"
                      /clone_lib="1008 - RescueMu Grid I"
                      /tissue_type="leaf"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Organ: leaf; Vector: RescueMu (engineered from
                      pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                      RescueMu is a 4.9 kb, modified maize Mu transposon
                      designed to allow plasmid rescue from total genomic DNA.
                      Mu elements insert preferentially into transcription
                      units. For more information on RescueMu, go to the  web
                      site www.zmdb.iastate.edu and follow the links for
                      'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                      extracted from leaf punches,  double digested using BamHI
                      and BglII, and ligated to form circular plasmids.  DH10B
                      cells were transformed and then screened on LB plates with
                      ampicillin."
BASE COUNT       22 a     19 c     22 g     10 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 17;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 8.7e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         60 AAGACGATGACGACGA 75
              ||||||||||||||||
Db          1 AAGACGATGACGACGA 16


RESULT 13
BJ034007
LOCUS       BJ034007                  79 bp    mRNA    linear   EST 05-DEC-2001
DEFINITION  BJ034007 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL025o10 5', mRNA sequence.
ACCESSION   BJ034007
VERSION     BJ034007.1  GI:17379716
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .79
                      /organism="Xenopus laevis"
                      /db_xref="taxon:8355"
                      /clone="XL025o10"
                      /clone_lib="NIBB Mochii normalized Xenopus neurula
                      library"
                      /tissue_type="whole embryo"
                      /dev_stage="stage 15"
                      /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                      were oligo-dT primed and directionally cloned. Staging
                      according to  Nieuwkoop and Faber. Library is subtracted
                      and was constructed by N. Garrett  and A.M. Zorn,
                      (Wellcome/CRC Institute). "
BASE COUNT       17 a     14 c     25 g     23 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 13;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 8.9e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 TCTGAGGGGACCTCAG 95
              ||||||||||||||||
Db         55 TCTGAGGGGACCTCAG 70


RESULT 14
AA761095/c
LOCUS       AA761095                  82 bp    mRNA    linear   EST 07-FEB-1998
DEFINITION  ny13h05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1271673 3'
```

```
            similar to SW:GLI4_HUMAN P10075 GLI4 PROTEIN ;, mRNA sequence.
ACCESSION   AA761095
VERSION     AA761095.1  GI:2810025
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1924   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .82
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1271673"
                      /clone_lib="NCI_CGAP_GCB1"
                      /tissue_type="germinal center B cell"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was prepared from human tonsillar cells enriched for
                      germinal center B cells by flow sorting (CD20+, IgD-),
                      provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                      (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                      primed with a Not I - oligo(dT) primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-3'
                      ]. Double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization, and was
                      constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT        7 a     35 c     22 g     18 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 9e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         96 GGGCTGCCGAGCTGGG 111
              ||||||||||||||||
Db         69 GGGCTGCCGAGCTGGG 54


RESULT 15
AA620617/c
LOCUS       AA620617                 88 bp    mRNA    linear   EST 14-OCT-1997
DEFINITION  af84b06.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1048691
            3' similar to SW:KR4_HUMAN P10075 HKR4 PROTEIN ;, mRNA sequence.
ACCESSION   AA620617
VERSION     AA620617.1  GI:2524556
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1048691"
                      /clone_lib="Soares_testis_NHT"
                      /sex="male"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was prepared from mRNA obtained from Clontech Laboratories
                      , Inc., and primed with  a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                      Double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. Library
                      went through one round of normalization to Cot5, and was
                      constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       17 a     33 c     25 g     13 t
ORIGIN

  Query Match             0.8%;  Score 16;  DB 9;  Length 88;
  Best Local Similarity   100.0%;  Pred. No. 9.1e+03;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         96 GGGCTGCCGAGCTGGG 111
              ||||||||||||||||
Db         75 GGGCTGCCGAGCTGGG 60


Search completed: July  4, 2003, 10:37:47
Job time : 2908 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 4, 2003, 06:57:16 ; Search time 98 Seconds
(without alignments)
6305.650 Million cell updates/sec

Title: US-10-007-010-3
Perfect score: 2015
Sequence: 1 cggaggcacggaagatgagg.....atataaatgcaagtcttacg 2015

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18	0.9	20	2	US-08-910-629A-31	Sequence 31, Appl
	2	18	0.9	20	2	US-08-910-629A-42	Sequence 42, Appl
c	3	18	0.9	20	3	US-09-209-668-7	Sequence 7, Appli
c	4	18	0.9	20	3	US-09-287-796-31	Sequence 31, Appl
	5	18	0.9	20	3	US-09-287-796-42	Sequence 42, Appl
c	6	18	0.9	20	4	US-09-130-616-31	Sequence 31, Appl
	7	18	0.9	20	4	US-09-130-616-42	Sequence 42, Appl
c	8	17	0.8	20	2	US-08-730-876-2	Sequence 2, Appli
c	9	17	0.8	20	4	US-09-490-692-71	Sequence 71, Appl
c	10	17	0.8	23	1	US-08-222-616-2	Sequence 2, Appli
c	11	17	0.8	23	4	US-08-446-648-2	Sequence 2, Appli
c	12	17	0.8	23	5	PCT-US95-04228-2	Sequence 2, Appli
c	13	16	0.8	20	4	US-09-506-073-82	Sequence 82, Appl
	14	16	0.8	24	2	US-08-859-998-598	Sequence 598, App
	15	16	0.8	24	4	US-09-225-928-598	Sequence 598, App
c	16	15	0.7	18	3	US-08-951-923-51	Sequence 51, Appl
c	17	15	0.7	18	4	US-08-584-040-6218	Sequence 6218, Ap
	18	15	0.7	19	1	US-08-400-580A-11	Sequence 11, Appl
	19	15	0.7	31	2	US-08-942-423-51	Sequence 51, Appl
	20	15	0.7	36	3	US-08-951-923-52	Sequence 52, Appl
c	21	15	0.7	36	3	US-08-724-586-3	Sequence 3, Appli
c	22	15	0.7	36	4	US-09-421-632-3	Sequence 3, Appli
c	23	15	0.7	36	4	US-09-932-190-3	Sequence 3, Appli
c	24	15	0.7	45	2	US-08-039-198B-3	Sequence 3, Appli
	25	15	0.7	72	2	US-08-707-237A-47	Sequence 47, Appl
	26	14	0.7	17	4	US-08-584-040-7661	Sequence 7661, Ap
c	27	14	0.7	18	1	US-08-105-483-197	Sequence 197, App
c	28	14	0.7	18	1	US-08-220-151-78	Sequence 78, Appl
c	29	14	0.7	18	1	US-08-413-118-78	Sequence 78, Appl
c	30	14	0.7	18	1	US-08-224-657-54	Sequence 54, Appl
c	31	14	0.7	18	1	US-08-709-209-197	Sequence 197, App
c	32	14	0.7	18	1	US-08-458-101-197	Sequence 197, App
c	33	14	0.7	18	2	US-08-184-009-52	Sequence 52, Appl
	34	14	0.7	18	2	US-08-173-489C-11	Sequence 11, Appl
c	35	14	0.7	18	2	US-08-417-210A-52	Sequence 52, Appl
c	36	14	0.7	18	2	US-08-585-684B-2737	Sequence 2737, Ap
c	37	14	0.7	18	2	US-08-458-356-52	Sequence 52, Appl
c	38	14	0.7	18	3	US-08-473-446-78	Sequence 78, Appl
c	39	14	0.7	18	4	US-09-038-073-2737	Sequence 2737, Ap
c	40	14	0.7	18	4	US-08-460-736-52	Sequence 52, Appl
c	41	14	0.7	18	4	US-09-354-138-54	Sequence 54, Appl
	42	14	0.7	20	1	US-08-639-763-6	Sequence 6, Appli
c	43	14	0.7	20	4	US-09-270-542-161	Sequence 161, App
c	44	14	0.7	20	4	US-09-798-096-87	Sequence 87, Appl
c	45	14	0.7	21	1	US-08-056-200-44	Sequence 44, Appl

ALIGNMENTS

```
RESULT 1
US-08-910-629A-31/c
; Sequence 31, Application US/08910629A
; Patent No. 5877309
;  GENERAL INFORMATION:
;    APPLICANT:  Robert A. McKay
;    APPLICANT:  Nicholas M. Dean
;    APPLICANT:  Brett Monia
;    TITLE OF INVENTION:  ANTISENSE OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR THE MODULATION OF JNK
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Jane Massey Licata
;      STREET:  66 East Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 MB
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  PENTIUM
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,629A
;      FILING DATE:  August 13, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  Yes
US-08-910-629A-31

  Query Match             0.9%;  Score 18;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
```

```
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 2
US-08-910-629A-42
; Sequence 42, Application US/08910629A
; Patent No. 5877309
;  GENERAL INFORMATION:
;    APPLICANT:  Robert A. McKay
;    APPLICANT:  Nicholas M. Dean
;    APPLICANT:  Brett Monia
;    TITLE OF INVENTION:  ANTISENSE OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  COMPOSITONS AND METHODS FOR THE MODULATION OF JNK
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Jane Massey Licata
;      STREET:  66 East Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 MB
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  PENTIUM
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,629A
;      FILING DATE:  August 13, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
US-08-910-629A-42

  Query Match             0.9%;  Score 18;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db          1 GACTTTGGCCTGGCCCGG 18


RESULT 3
US-09-209-668-7/c
; Sequence 7, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
;  APPLICANT: Monia, Brett P.
;  APPLICANT: Xu, Xiaoxing S.
;  TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
;  TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
;  FILE REFERENCE: ISPH-0336
;  CURRENT APPLICATION NUMBER: US/09/209,668A
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: antisense sequence
US-09-209-668-7

  Query Match             0.9%;  Score 18;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 4
US-09-287-796-31/c
; Sequence 31, Application US/09287796A
; Patent No. 6133246
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0350
;  CURRENT APPLICATION NUMBER: US/09/287,796A
;  CURRENT FILING DATE: 1999-04-07
;  EARLIER APPLICATION NUMBER: 09/130,616
;  EARLIER FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 31
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-287-796-31

  Query Match             0.9%;  Score 18;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 5
US-09-287-796-42
; Sequence 42, Application US/09287796A
; Patent No. 6133246
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0350
;  CURRENT APPLICATION NUMBER: US/09/287,796A
```

```
;  CURRENT FILING DATE: 1999-04-07
;  EARLIER APPLICATION NUMBER: 09/130,616
;  EARLIER FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 42
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-287-796-42

  Query Match             0.9%;  Score 18;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db          1 GACTTTGGCCTGGCCCGG 18


RESULT 6
US-09-130-616-31/c
; Sequence 31, Application US/09130616C
; Patent No. 6221850
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0318
;  CURRENT APPLICATION NUMBER: US/09/130,616C
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 31
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-130-616-31

  Query Match             0.9%;  Score 18;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db         20 GACTTTGGCCTGGCCCGG 3


RESULT 7
US-09-130-616-42
; Sequence 42, Application US/09130616C
; Patent No. 6221850
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0318
;  CURRENT APPLICATION NUMBER: US/09/130,616C
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 42
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-130-616-42

  Query Match             0.9%;  Score 18;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1300 GACTTTGGCCTGGCCCGG 1317
              ||||||||||||||||||
Db          1 GACTTTGGCCTGGCCCGG 18


RESULT 8
US-08-730-876-2/c
; Sequence 2, Application US/08730876
; Patent No. 5859314
;  GENERAL INFORMATION:
;    APPLICANT:  HIBBS, Margaret L.;
;    APPLICANT:  DUNN, Ashley R.;
;    APPLICANT:  GRAILL, Dianne;
;    APPLICANT:  HODGSON George;
;    APPLICANT:  TARLINGTON, David M.;
;    APPLICANT:  ARMES, Jane
;    TITLE OF INVENTION:  ANIMALS WITH TARGETED GENE DELETION
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44mb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/730,876
;      FILING DATE:  18-Oct-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/005,578
;      FILING DATE:  20-Oct-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5859314man D. Hanson
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5369 - JEL/NDH/SLH
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-730-876-2

  Query Match             0.8%;  Score 17;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        916 GGGCAGTTTGGGGAAGT 932
              |||||||||||||||||
```

```
Db         17 GGGCAGTTTGGGGAAGT 1


RESULT 9
US-09-490-692-71/c
; Sequence 71, Application US/09490692
; Patent No. 6180353
; GENERAL INFORMATION:
;  APPLICANT: Nicholas M. Dean
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF DAXX EXPRESSION
;  FILE REFERENCE: RTS-0120
;  CURRENT APPLICATION NUMBER: US/09/490,692
;  CURRENT FILING DATE: 2000-01-24
;  NUMBER OF SEQ ID NOS: 176
; SEQ ID NO 71
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-490-692-71

  Query Match             0.8%;  Score 17;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 TCAGGAGGATGATGAAG 44
              |||||||||||||||||
Db         18 TCAGGAGGATGATGAAG 2


RESULT 10
US-08-222-616-2/c
; Sequence 2, Application US/08222616
; Patent No. 5635177
;  GENERAL INFORMATION:
;    APPLICANT:  Bennett, Brian D.
;    APPLICANT:  Goeddel, David
;    APPLICANT:  Lee, James M.
;    APPLICANT:  Matthews, William
;    APPLICANT:  Tsai, Siao Ping
;    APPLICANT:  Wood, William I.
;    TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST
;    TITLE OF INVENTION:  ANTIBODIES
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,616
;      FILING DATE:  4-APR-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/00586
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/826935
;      FILING DATE:  22-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:  821P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-222-616-2

  Query Match             0.8%;  Score 17;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 11
US-08-446-648-2/c
; Sequence 2, Application US/08446648
; Patent No. 6331302
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  Bennett, Brian D.
;    APPLICANT:  Goeddel, David
;    APPLICANT:  Lee, James M.
;    APPLICANT:  Matthews, William
;    APPLICANT:  Tsai, Siao Ping
;    APPLICANT:  Wood, William I.
;    TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,648
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/222616
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  40,378
;      REFERENCE/DOCKET NUMBER:  P0821P3PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-446-648-2

  Query Match             0.8%;  Score 17;  DB 4;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 12
PCT-US95-04228-2/c
; Sequence 2, Application PC/TUS9504228
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  Bennett, Brian D.
;    APPLICANT:  Goeddel, David
;    APPLICANT:  Lee, James M.
;    APPLICANT:  Matthews, William
;    APPLICANT:  Tsai, Siao Ping
;    APPLICANT:  Wood, William I.
;    TITLE OF INVENTION:  PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04228
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/222616
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wendy M. Lee
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  821P3PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US95-04228-2

  Query Match             0.8%;  Score 17;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1420 GACGTCTGGTCCTTTGG 1436
              |||||||||||||||||
Db         23 GACGTCTGGTCCTTTGG 7


RESULT 13
US-09-506-073-82/c
; Sequence 82, Application US/09506073
; Patent No. 6410518
; GENERAL INFORMATION:
;  APPLICANT: Monia, Brett P.
;  TITLE OF INVENTION: Antisense Oligonucleotide Modulation of raf Gene Expression
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/506,073
;  CURRENT FILING DATE: 2000-02-18
;  EARLIER APPLICATION NUMBER: US 09/143,214
;  EARLIER FILING DATE: 1998-08-28
;  EARLIER APPLICATION NUMBER: PCT/US98/13961
;  EARLIER FILING DATE: 1998-07-06
;  EARLIER APPLICATION NUMBER: US 08/888,982
;  EARLIER FILING DATE: 1997-07-07
;  EARLIER APPLICATION NUMBER: US 08/756,806
;  EARLIER FILING DATE: 1996-11-26
;  EARLIER APPLICATION NUMBER: PCT/US95/07111
;  EARLIER FILING DATE: 1995-05-31
;  EARLIER APPLICATION NUMBER: US 08/250,856
;  EARLIER FILING DATE: 1994-05-31
;  NUMBER OF SEQ ID NOS: 130
; SEQ ID NO 82
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: antisense sequence
US-09-506-073-82

  Query Match             0.8%;  Score 16;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       155 GAGCGGGCGCCAGGAT 170
             ||||||||||||||||
Db        20 GAGCGGGCGCCAGGAT 5


RESULT 14
US-08-859-998-598
; Sequence 598, Application US/08859998
; Patent No. 5994076
;  GENERAL INFORMATION:
;    APPLICANT:  Chenchik, Alex
;    APPLICANT:  Jokhadze, George
;    APPLICANT:  Bibilashvilli, Robert
;    TITLE OF INVENTION:  METHOD OF ASSAYING DIFFERENTIAL
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1375
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  2200 Sand Hill Road, Suite 100
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/859,998
;      FILING DATE:  21-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Field, Bret E.
;      REGISTRATION NUMBER:  37,620
;      REFERENCE/DOCKET NUMBER:  09096/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  598:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
```

```
;    FEATURE:
;      OTHER INFORMATION:  oligonucleotide primer
US-08-859-998-598

  Query Match             0.8%;  Score 16;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1236 CATCCACCGAGACCTC 1251
              ||||||||||||||||
Db          8 CATCCACCGAGACCTC 23


RESULT 15
US-09-225-928-598
; Sequence 598, Application US/09225928
; Patent No. 6352829
;   GENERAL INFORMATION:
;        APPLICANT: Chenchik, Alex
;                   Jokhadze, George
;                   Bibilashvilli, Robert
;        TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
;                            EXPRESSION
;        NUMBER OF SEQUENCES: 1375
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson, P.C.
;             STREET: 2200 Sand Hill Road, Suite 100
;             CITY: Menlo Park
;             STATE: CA
;             COUNTRY: US
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/225,928
;             FILING DATE: 05-Jan-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/859,998
;             FILING DATE: 21-MAY-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Field, Bret E.
;             REGISTRATION NUMBER: 37,620
;             REFERENCE/DOCKET NUMBER: 09096/002001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-322-5070
;             TELEFAX: 415-854-0875
;   INFORMATION FOR SEQ ID NO: 598:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 24 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        FEATURE:
;             OTHER INFORMATION:  oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 598:
US-09-225-928-598

  Query Match             0.8%;  Score 16;  DB 4;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1236 CATCCACCGAGACCTC 1251
              ||||||||||||||||
Db          8 CATCCACCGAGACCTC 23


Search completed: July  4, 2003, 10:39:34
Job time : 100 secs
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 TCAGGAGGATGATGAAG 44
              |||||||||||||||||
Db         18 TCAGGAGGATGATGAAG 2


RESULT 15
AX201544/c
LOCUS       AX201544                  21 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 223 from Patent WO0153486.
ACCESSION   AX201544
VERSION     AX201544.1  GI:15391386
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
            Stone,D.M., Watanabe,C.K. and Wood,W.I.
  TITLE     Compositions and methods for the treatment of tumour
  JOURNAL   Patent: WO 0153486-A 223 26-JUL-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide Probe."
BASE COUNT        7 a      8 c      3 g      3 t
ORIGIN

  Query Match           0.8%;  Score 17;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        907 CTTGGAGCTGGGCAGTT 923
              |||||||||||||||||
Db         19 CTTGGAGCTGGGCAGTT 3


Search completed: July  4, 2003, 08:32:26
Job time : 5240 secs
```